

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 20:50:16 ; Search time 2114 Seconds
(without alignments)
6991.153 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399

Perfect score: 260
Sequence: 1 gagttggcgaccatggtgg.....cccgcttggaataaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_in:*

3: gb_env:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pr:*

9: gb_ro:*

10: gb_sta:*

11: gb_sy:*

12: gb_un:*

13: gb_vi:*

14: gb_hgt:*

15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	9399	6	AX057392 Sequence
2	260	100.0	9399	11	AY243572 Synthetic
3	260	100.0	9399	13	AF179612 Hepatitis
4	258.4	99.4	309	6	AR403987 Sequence
5	258.4	99.4	309	6	AX055777 Sequence
6	258.4	99.4	357	13	HGB18973
7	258.4	99.4	8027	13	GV1428955
8	258.4	99.4	8069	6	AX805211
9	258.4	99.4	9397	6	AX805212
10	258.4	99.4	9397	13	HG8277947
11	257.4	99.0	259	6	AR403988
12	257.4	99.0	259	6	AR403989
13	257.4	99.0	259	6	AR403990
14	257.4	99.0	259	6	AR403991
15	257.4	99.0	259	6	AX055778
16	257.4	99.0	259	6	AX055779
17	257.4	99.0	259	6	AX055780
18	257.4	99.0	259	6	AX055781

19	82	31.5	82	6	AR404000	AR404000 Sequence
20	82	31.5	82	6	AX055790	AX055790 Sequence
21	34.8	13.4	229312	9	AC134603	AC134603 Mus muscu
22	34.6	13.3	204413	9	AL645910	AL645910 Mouse DNA
23	33.2	12.8	110000	1	AE017283_19	Continuation (20 o
24	33.2	12.8	182634	8	AC142353	AC142353 Pan trogl
25	33.2	12.8	185516	8	AC004485	AC004485 Homo sapi
26	32.8	12.6	171578	14	AC149190	AC149190 Papio anu
27	32.8	12.6	180140	14	AC145521	AC145521 Papio ham
28	32.8	12.6	250175	14	AC160494	AC160494 Bos tauru
29	32.6	12.5	172715	9	CNS07EGP	ALS91826 BAC 13C18
30	32.6	12.5	219709	14	AC156237	AC156237 Bos tauru
31	32.4	12.5	2354	13	PEM76VP1	D00056 Monkey B-ly
32	32.4	12.5	2354	13	PEM76VP1	M14494 Monkey B-ly
33	32.4	12.5	160738	14	AC161787	AC161787 Cercopithec
34	32.2	12.4	3003	6	AX780548	AX780548 Sequence
35	32.2	12.4	82414	14	AC129265	AC129265 Rattus no
36	32.2	12.4	185273	14	AC019038	AC019038 Homo sapi
37	32.2	12.4	198912	9	AC132380	AC132380 Mus muscu
38	32.2	12.4	218129	14	AC136560	AC136560 Rattus no
39	32.2	12.4	219339	14	AC121005	AC121005 Rattus no
40	32.2	12.4	234545	14	AC106571	AC106571 Rattus no
41	32.2	12.4	237041	14	AC133483	AC133483 Rattus no
42	32	12.3	3179	6	AX179748	AX179748 Sequence
43	32	12.3	67699	15	AB016815_8	Continuation (9 of
44	32	12.3	154622	14	AC126336	AC126336 Homo sapi
45	32	12.3	178714	14	AC129504	AC129504 Homo sapi

ALIGNMENTS

RESULT 1
AX057392
LOCUS AX057392 9399 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 1 from Patent WO0075337.
ACCESSION AX057392
VERSION AX057392.1 GI:12310132
KEYWORDS
SOURCE Hepatitis GB virus B
ORGANISM Hepatitis GB virus B
Virus; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE 1
AUTHORS Bukh, J., Yanagi, M., Emerson, S.U. and Purcell, R.H.
TITLE Infectious cdna clone of gb virus b and uses thereof
JOURNAL Patent: WO 0075337-A 1 14-DEC-2000;
THE GOVERNMENT OF THE UNITED STATES OF AMERICA (US)
FEATURES
Location/Qualifiers
source 1..9399
/organism="Hepatitis GB virus B"
/mol_type="unassigned DNA"
/db_xref="taxon:39113"

ORIGIN

Query Match 100.0%; Score 260; DB 6; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1.7e-67;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9140 GAGTTTGGGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9199
QY 61 GAGTCCCTTCCTGGCTATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTG 120
DB 9200 GAGTCCCTTCCTGGCTATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTG 9259
QY 121 TGGGAAGCAGTCAGTATAATTCCTCGTGTGTGTGTCAGCCCTCAGACGATTTTGTCCG 180
DB 9260 TGGGAAGCAGTCAGTATAATTCCTCGTGTGTGTGTCAGCCCTCAGACGATTTTGTCCG 9319
QY 181 CTGTGTCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTGTTCGAAGCGAGGCAACC 240
DB 9320 CTGTGTCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTGTTCGAAGCGAGGCAACC 9379

/organism="Hepatitis GB virus B"
/mol_type="genomic RNA"
/specific_host="unknown"
/db_xref="taxon:39113"
/note="most closely related to hepatitis C virus;
experimental infection of tamarins causes acute resolving
hepatitis"
446..9040
/codon_start=1
/product="polyprotein"
/protein_id="AA01368.1"
/db_xref="GI:6014505"

CDS

/translation="MPVISTQTSVPAPRTRKNNKQTSQSPVSIKTSVERQORAKRV
QRDAPRNYKTAGHDGQTLQAALPAHGWGRDPRHKSRLNGLILLPLGWTGDVT
THPLVAGVAVRQVQVIRLEEDGNWATGFWGHLFVCLLSLACSCGARVT
DPDNTTILTCORNOVIYCSPTSLGPCVICADECWPANPYISHPNCTGTDS
FLADHIDFVMGALVTCDDALDGLGACVCLVQDMLVRHLIHLIDNETGTCYLEVPTG
IDPGLFGIWMAGKVEAVIELTKLASQVPAIATMFSVHYLAVALIYYASRGKY
QLLLALMIYIATSNPIRVPTGSIABFCSPLMPCPCSHSYLSENVSEVICSPKWT
RPTILEYNNISWYPTTPGARGCMKFNNTGCCRIINVPSTYTMGTDAWMDTRN
TYEACGVPTMLTTHWNGSALKLAILOYPGSKEMKPHNMSGHLYPFGSDTPIVIFY
DPVNSTLPPRWRALPGTPVWRGSLQVPGPYSDVKLATGLITKDKAKNYOVL
YSAGALSITGVTTKAVLILGLGSKYLILAVLCYSLCFCGRASGYLPRVLPSSQ
YLOAGDVLKSAQVAPFALIFPICCYLCRLRYAALLGFVPMAGLPLTFVFAAAAQ
PDYDMMVLLVAGLVLMAGNRGHRIALLVGPVPLVALLTLHLVTPASAFDETLIG
LTIPTVALVMSRFGFFFAHLLPRCALVNSYLMQRWENFWNVTLRPERFLVLVCFP
GATYDALVTFCVCHVALCLTSSAASFCTDSRVRAHMLVRLGKCHAWYSHYLVKFF
LLVFCENGVEFYKHLHGVLPNDPASKLPLOEPPPPRPGKARVYRNEGRRLACGTVD
GLPVVARGDLVFLAGLAMPDGAITAPFTLQCLSERGTLSMAVWMTGIDPRWTG
IFRUGSLATSYMGVCDNVLITAHGSGRRRLAHPGTSIHPIFTVDAANDQDIYQPCG
AGSLTRCGTGBGTGLVRLGSLVEVNSDDPYCVCALPMNAVAKSGGAPILCSGG
HVIWFTASRSGGVSQIRVRPLVCAQHPQVTAHATLDKTPVNEYSQIILAPT
GSGKTKPLSPYMOBKLYVLNPSVATASMPKMYHATYGVNPNYFNGKCTNGAS
LTSYTYGVALTGACSRNDVLIICDECHADTATTVLIGIKVLEAPSKNVLVLATAT
PGVPTPHANTTELQIDBGTIPHGKKIKEENLXKXERHLIPEATKXCHDELANELA
RGTGITSYRGCDLSIKIPEGDCVVVADALCTGYIGDFSDVDCSLMVEGTCHVDLD
PTFTMGVRCVGSIAIVKQGRGRTGRGRIYIYVDSCTPSGMVPECNIVEAFDAAD
AWYGLSSTAQTLIDTYRTQPLPAIGANLDEWADLFSMVNPPEFVNTAKRTADNYV
LLTAQLOLCHOYGAAVNDAPRWQARGLKPKCGVLRWDGACADCPGPSEVTRYQ
MCFTEVNTSGTAAALVAGVGMAYLAIDTFGATCYRRCSITSVPTGATVAPVUDEE
IYBECASIPLEAMVAADIKLSTITTSPTFLETALSKLNTFLGPHATLAIIEYC
CGLVLPNPFASCVFAGITTPPHKIKMFLSFGGATSKATDARGALAFMAG
AAGTALGTSTWYFVDFMLGGYAAASCTCLTKMGEMWTDQLAGLVYSAFNPA
GVVGLSACAFALMTAGDPHNPRLTLMLARSNTVCEYFIATDIRKILGLEAS
TPWYSACIRIHTPTBDDGLIAGLEIWOYCNFVFCVNLKAGVQSNVNIPOG
PEYSCQYKGPWIGSMLOARCPGABLI PSVENGFAKLYKGPCTCSNVRGAVPN
AKLCSARPDPFDITSLVYNGVDTYCKFERMGDHI FTVATSSPNVCTQVPTLRAA
VAVDGVQCYLGEPEKTPMTTACCYPGPGKTKVLPFRVDGHTPGVRMLNLRDAL
ETNDCNNTNTPSDBAASVAFVQELRRTNQLLEAI SAGVDITTKLPAPSEEVVVRK
RQFRARTGSLTLPVPPRSPVSPVCPESLQSPDLEPSNLPPSPVLOLAMPPLLGA
GECNPFTAI GAWTETGGPDLPSPYPPKKEVSEWSDSSSTATTASSVVTGPPYPKI
RGKISTQAPAKRPTKKLKGSEFSCSMSTWTDVISPKTASKVLSATRAITSGFLQ
RSLVTPTEPRDLRKRQVITNRQPLFPSPYHKQVRLAKERKAKVGVWMDYDEVAH
TPSKSAKSHITGLRDTVRSGAKAVLDLQKCEAGEIPSHYRTVIVPKEEVPKT
POKPTKPRLIISPHLEMRVCKMYQGVAPDVVKAVMGDAGVDFDPRTRKLLSM
WSPDVGATCDTVCFDSTITPEDIWETDIYSAAKLSQDHRAGIHTIARQLYAGGPMI
AYDREIGVRCRSGSVTSSNSLITCWLKYNAAEAOAKNPRFLICGDDCTVIWK
SAGADARQMRVFPASWVKWGPQDCVPQPKYLEELTSCSNVTSIGTSGKPYE
LTRDPRIPLGRSGELGNPSAAMWIGLIHYPCLWVSRVLAVHFMQMLFEDKPE
TYTFPMYGNVTPVEDLPSIIAGVHLEFNSVRYTNAELRVQSUTDMTPPLRA
WRKKARAVLASAGRRGGAAKLARFLWHATSRPLDLDKTSVARYTTTFNYCDVTSPE
GDVFTTPQRRLQKLVKVLAVIVFALGLIAVGLAIS"

ORIGIN

Query Match 100.0%; Score 260; DB 13; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1.7e-67;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACATGGTCAGAACCGTTTCGGGTGAGGAGTCTCTGGCTGTG 60
DB 9140 GAGTTTGGCGACATGGTCAGAACCGTTTCGGGTGAGGAGTCTCTGGCTGTG 9199
QY 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 120

Db 9200 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 9259
QY 121 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACCGCTCAGCAGCTATTGTTGTCG 180
Db 9260 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACCGCTCAGCAGCTATTGTTGTCG 9319
QY 181 CTGTGCAGAGCGTAGTACCAAGGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
Db 9320 CTGTGCAGAGCGTAGTACCAAGGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
QY 241 CCCGCTTGGAAATTAATAACT 260
Db 9380 CCCGCTTGGAAATTAATAACT 9399
RESULT 4
AR403987 AR403987 309 bp DNA linear PAT 18-DEC-2003
LOCUS Sequence 1 from patent US 6627437.
DEFINITION AR403987
ACCESSION AR403987
VERSION AR403987.1 GI:40151915
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 309)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 1 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;
FEATURES
source 1..309
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.4%; Score 258.4; DB 6; Length 309;
Best Local Similarity 99.6%; Pred. No. 8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAGGAGTCTCTGGCTGTG 60
Db 50 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAGGAGTCTCTGGCTGTG 109
QY 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 120
Db 110 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 169
QY 121 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACCGCTCAGCAGCTATTGTTGTCG 180
Db 170 TGGGAAGCAGTCAGTATAATTTCCCGTCGTGTGTGTGACCGCTCAGCAGCTATTGTTGTCG 229
QY 181 CTGTGCAGAGCGTAGTACCAAGGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
Db 230 CTGTGCAGAGCGTAGTACCAAGGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 289
QY 241 CCCGCTTGGAAATTAATAACT 260
Db 290 CCCGCTTGGAAATTAATAACT 309
RESULT 5
AX055777
LOCUS Sequence 1 from Patent WO0073466.
DEFINITION AX055777
ACCESSION AX055777
VERSION AX055777.1 GI:12228889
KEYWORDS
SOURCE unidentified
ORGANISM unidentified

```
unclassified sequences.
REFERENCE
1
AUTHORS Traboni,C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 1 07-DEC-2000;
        ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI, S.P.A.
        (IT)
FEATURES
    source      Location/Qualifiers
                1..309
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
                /note="GBV-B-like virus"
                1..309
                /note="cDNA complementary to the last 49 published 3'UTR
                nucleotides of GBV-B plus a novel nucleotide (C) in
                position 44 and 259 novel nucleotide sequence at the 3',
                end."
3'UTR
ORIGIN
Query Match      99.4%; Score 258.4; DB 6; Length 309;
Best Local Similarity 99.6%; Pred. No. 8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 60
    |||
Db 50 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 109
    |||
Qy 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGAAGCCATGCTCTGAAGGGGAT 120
    |||
Db 110 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGAAGCCATGCTCTGAAGGGGAT 169
    |||
Qy 121 TGGGAAGCAGTCAGTATTAATCCCTCGTGTGTGTGACGCTCAGACGTAATTTGTCCG 180
    |||
Db 170 TGGGAAGCAGTCAGTATTAATCCCTCGTGTGTGTGACGCTCAGACGTAATTTGTCCG 229
    |||
Qy 181 CTGTCAGAGCTAGTACCAAGGCTGCACCCCGGTTTGTTCGAAGGGGCAACC 240
    |||
Db 230 CTGTCAGAGCTAGTACCAAGGCTGCACCCCGGTTTGTTCGAAGGGGCAACC 289
    |||
Qy 241 CCCGCTTGGAAATTAATAACT 260
    |||
Db 290 CCCGCTTGGAAATTAATAACT 309
    |||
RESULT 6
HGB18973
LOCUS HGB18973 357 bp RNA linear VRL 19-NOV-1999
DEFINITION Hepatitis GB virus B partial 3'UTR region.
ACCESSION Y18973
VERSION Y18973.1 GI:6018427
KEYWORDS 3' UTR.
SOURCE Hepatitis GB virus B
ORGANISM Hepatitis GB virus B
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE
1
AUTHORS Sbardellati,A., Scarselli,E., Tomei,L., Kekule,A.S. and Traboni,C.
TITLE Identification of a novel sequence at the 3' end of the GB virus B
genome
JOURNAL J. Virol. 73 (12), 10546-10550 (1999)
PUBMED 10559376
REFERENCE
2 (bases 1 to 357)
AUTHORS Traboni,C.
TITLE Direct Submission
JOURNAL Submitted (29-APR-1999) C. Traboni, IRBM P. Angeletti, Ist. di
        Ricerche di Biologia Molecolare, Via Pontina. Km. 30.600, 00040
        Pomezia, Rome, ITALY
FEATURES
    source      Location/Qualifiers
                1..357
                /organism="Hepatitis GB virus B"
                /mol_type="genomic RNA"
                /specific host="tamarin"
                /db_xref="taxon:39113"
                .
                .
                .
unclassified sequences.
REFERENCE
1
AUTHORS Traboni,C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 1 07-DEC-2000;
        ISTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI, S.P.A.
        (IT)
FEATURES
    source      Location/Qualifiers
                1..357
                /organism="unidentified"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32644"
                /note="GBV-B-like virus"
                1..357
                /note="cDNA complementary to the last 49 published 3'UTR
                nucleotides of GBV-B plus a novel nucleotide (C) in
                position 44 and 259 novel nucleotide sequence at the 3',
                end."
3'UTR
ORIGIN
Query Match      99.4%; Score 258.4; DB 13; Length 357;
Best Local Similarity 99.6%; Pred. No. 7.8e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 60
    |||
Db 98 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 157
    |||
Qy 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGAAGCCATGCTCTGGCTGTG 120
    |||
Db 158 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGAAGCCATGCTCTGGCTGTG 217
    |||
Qy 121 TGGGAAGCAGTCAGTATTAATCCCTCGTGTGTGTGACGCTCAGACGTAATTTGTCCG 180
    |||
Db 218 TGGGAAGCAGTCAGTATTAATCCCTCGTGTGTGTGACGCTCAGACGTAATTTGTCCG 277
    |||
Qy 181 CTGTCAGAGCTAGTACCAAGGCTGCACCCCGGTTTGTTCGAAGGGGCAACC 240
    |||
Db 278 CTGTCAGAGCTAGTACCAAGGCTGCACCCCGGTTTGTTCGAAGGGGCAACC 337
    |||
Qy 241 CCCGCTTGGAAATTAATAACT 260
    |||
Db 338 CCCGCTTGGAAATTAATAACT 357
    |||
RESULT 7
GV1428955
LOCUS GV1428955 8027 bp mRNA linear VRL 15-APR-2005
DEFINITION Hepatitis GB virus B subgenomic replicon neoRepB.
ACCESSION AJ428955
VERSION AJ428955.1 GI:21272785
KEYWORDS core-neo fusion protein; core-neo gene; polyprotein.
SOURCE Hepatitis GB virus B
ORGANISM Hepatitis GB virus B
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Flaviviridae.
REFERENCE
1
AUTHORS De Tomassi,A., Pizzuti,M., Graziani,R., Sbardellati,A.,
        Altamura,S., Paonessa,G. and Traboni,C.
TITLE Cell clones selected from the Huh7 human hepatoma cell line support
        efficient replication of a subgenomic GB virus B replicon
JOURNAL J. Virol. 76 (15), 7736-7746 (2002)
PUBMED 12097587
REFERENCE
2 (bases 1 to 8027)
AUTHORS Traboni,C.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-2002) Traboni C., Biochemistry, IRBM P. Angeletti,
        Via Pontina, Km.30, 600. 00040 Pomezia (Roma), ITALY
COMMENT related sequence AJ27947.
FEATURES
    source      Location/Qualifiers
                join(1..1281,1893..8027)
                /organism="Hepatitis GB virus B"
                /mol_type="mRNA"
                /isolate="FL3"
                /db_xref="taxon:39113"
                /focus
                1282..1892
                /organism="Encephalomyocarditis virus"
                /mol_type="mRNA"
                /db_xref="taxon:12104"
                1..445
                /evidence=experimental
                446..1273
                /function="core-neo fusion protein"
                .
                .
                .
```


ORIGIN

Query Match 99.4%; Score 258.4; DB 13; Length 9397;
Best Local Similarity 99.6%; Pred. No. 5.1e-67;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 60
DB 9138 GAGTTTGGCGACATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGAT 9197

QY 61 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGCTGAGGAGTCTCGCTGTG 120
DB 9198 GAGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGCTGAGGAGTCTCGCTGTG 9257

QY 121 TGGGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTCCG 180
DB 9258 TGGGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTCCG 9317

QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGACCCCGGTTTTTTGTTCAGCGGAGGCAACC 240
DB 9318 CTGTGCAGAGCGTAGTACCAAGGGCTGACCCCGGTTTTTTGTTCAGCGGAGGCAACC 9377

QY 241 CCGCTTGGAAATTAATAACT 260
DB 9378 CCGCTTGGAAATTAATAACT 9397

RESULT 11

AR403988
LOCUS AR403988 259 bp RNA linear PAT 18-DEC-2003
DEFINITION Sequence 2 from patent US 6627437.
ACCESSION AR403988
VERSION AR403988.1 GI:40151916
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 259)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 2 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;

FEATURES

source
1..259
/organism="unknown"
/mol_type="unassigned RNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.8%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATG 61
DB 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATG 60

QY 62 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCGCTGTGT 121
DB 61 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCGCTGTGT 120

QY 122 GGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTGCCG 181
DB 121 GGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTGCCG 180

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACC 241
DB 181 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACC 240

QY 242 CCGCTTGGAAATTAATAACT 260
DB 241 CCGCTTGGAAATTAATAACT 259

RESULT 12

AR403989/c
LOCUS AR403989 259 bp RNA linear PAT 18-DEC-2003
DEFINITION Sequence 3 from patent US 6627437.
ACCESSION AR403989
VERSION AR403989.1 GI:40151917
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 259)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 3 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;

FEATURES

source
1..259
/organism="unknown"
/mol_type="unassigned RNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATG 61
DB 259 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGCTCTGAAGGGGATG 200

QY 62 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCGCTGTGT 121
DB 199 ACCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCGCTGTGT 140

QY 122 GGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTGCCG 181
DB 139 GGAAGCAGTCAGTATATAATCCCGTCGTGTGTGACGCTCAGACGATATTGTGCCG 80

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACC 241
DB 79 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCAGCGGAGGCAACC 20

QY 242 CCGCTTGGAAATTAATAACT 260
DB 19 CCGCTTGGAAATTAATAACT 1

RESULT 13

AR403990
LOCUS AR403990 259 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 4 from patent US 6627437.
ACCESSION AR403990
VERSION AR403990.1 GI:40151918
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 259)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 4 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;

FEATURES

source
1..259
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
|||||

QY 62 AGTCCCTCTCGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 121
|||||
Db 61 AGTCCCTCTCGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 120
|||||

QY 122 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 181
|||||
Db 121 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 180
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 241
|||||
Db 181 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 241 CCGCTTGGAAATTAATAACT 259
|||||

RESULT 14
AR403991/c

LOCUS AR403991 259 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 5 from patent US 6627437.
ACCESSION AR403991
VERSION AR403991.1 GI:40151919
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 259)
AUTHORS Traboni, C.
TITLE GBV sequence
JOURNAL Patent: US 6627437-A 5 30-SEP-2003;
Istituto di Ricerche di Biologia Molecolare P. Angeletti S.p.A.;
Pomezia;
WOX;

FEATURES
Location/Qualifiers
source 1..259
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 259 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 200
|||||

QY 62 AGTCCCTCTCGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 121
|||||
Db 199 AGTCCCTCTCGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 140
|||||

QY 122 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 181
|||||
Db 139 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 80
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 241
|||||
Db 79 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 19 CCGCTTGGAAATTAATAACT 1
|||||

RESULT 15
AX055778

LOCUS AX055778 259 bp RNA linear PAT 13-JAN-2001
DEFINITION Sequence 2 from Patent WO0073466.
ACCESSION AX055778
VERSION AX055778.1 GI:12228890
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 unclassified sequences.
AUTHORS Traboni, C.I.
TITLE Novel gbv sequence
JOURNAL Patent: WO 0073466-A 2 07-DEC-2000;
ISITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI; S.P.A.
(IT)

FEATURES
Location/Qualifiers
source 1..259
/organism="unidentified"
/mol_type="unassigned RNA"
/db_xref="taxon:32644"
/note="GBV-B-like virus"

ORIGIN

Query Match 99.0%; Score 257.4; DB 6; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.6e-66;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
|||||
Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60
|||||

QY 62 AGTCCCTCTCGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 121
|||||
Db 61 AGTCCCTCTCGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGGAGTCTCGGCTGTGT 120
|||||

QY 122 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 181
|||||
Db 121 GGAAGACAGTCAAGTAATTCCTCGTGTGTGGTGACGCCCTCACGACGATTTTGTCCGC 180
|||||

QY 182 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 241
|||||
Db 181 TGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTCCAAAGCGAGGGCAACCC 240
|||||

QY 242 CCGCTTGGAAATTAATAACT 260
|||||
Db 241 CCGCTTGGAAATTAATAACT 259
|||||

Search completed: December 7, 2005, 21:43:47
Job time : 2117 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 19:45:21 ; Search time 366 Seconds

(without alignments)
4734.481 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399

Perfect score: 260

Sequence: 1 gaggttggcgaccatgtgg.....cccgttggaattaaaaact 260

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002bs.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260	100.0	9399	4	AAF23484
2	260	100.0	9399	12	ADJ56732
3	260	100.0	9399	12	ADJ64244
4	258.4	99.4	309	4	AAC92010
5	258.4	99.4	8069	9	ADA77751
6	258.4	99.4	9397	9	ADA77752
7	257.4	99.0	259	4	AAC91998
8	254.4	97.8	260	12	ADJ56731
9	254.4	97.8	260	12	ADJ64243
10	248.8	95.7	362	5	AAS15946
11	245.4	94.4	258	4	AAC91996
12	108	41.5	108	12	ADJ64259
13	82	31.5	82	4	AAC92012
14	66	25.4	108	12	ADJ56742
15	66	25.4	108	12	ADJ64254
16	32.2	12.4	3003	10	ADF82149
17	32	12.3	3179	4	AAD09560
18	31.2	12.0	306	6	ABN25772
19	31.2	12.0	174155	14	AEA07466

20	31	11.9	1095	6	ABS64612
21	30.6	11.8	3684	13	ADS96469
22	30.6	11.8	3772	4	ABL22533
23	30.4	11.7	522	3	AAC81741
24	30.4	11.7	2245	6	AAD33249
25	30.4	11.7	2342	3	AA93110
26	30.4	11.7	3229	8	ABZ36284
27	30.4	11.7	110000	14	ADZ12821_2
28	30.2	11.6	1012	4	AAS59679
29	30.2	11.6	1012	8	ACF64608
30	30.2	11.6	1308	4	AAS59670
31	30.2	11.6	1308	8	ACF64599
32	30	11.5	549	13	ADQ49573
33	30	11.5	2499	12	ADO07760
34	30	11.5	3855	4	ABL30105
35	30	11.5	4497	4	ABL05455
36	30	11.5	10907	4	ABL30104
37	30	11.5	13815	4	ABL05454
38	30	11.5	49999	2	AAZ23390
39	29.8	11.5	1411	13	ADX47507
40	29.8	11.5	24183	5	AAS21771
41	29.6	11.4	4659	14	ADY18412
42	29.4	11.3	460	12	ACH91975
43	29.4	11.3	1250	12	ADI24471
44	29.4	11.3	1306	10	ADJ80204
45	29.4	11.3	1497	10	ADF09669

ALIGNMENTS

RESULT 1

AAF23484

ID AAF23484 standard; DNA; 9399 BP.

XX AAF23484;

XX AAF23484;

DT 11-SEP-2003 (revised)

DT 21-MAR-2001 (first entry)

XX GBV-B virus genome.

DE GBV-B virus genome.

XX GBV-B; hepatitis C virus; HCV; vaccine; ds.

XX Hepatitis GB virus B.

XX WO200075337-A1.

XX 14-DEC-2000.

XX 02-JUN-2000; 2000WO-US015293.

XX 04-JUN-1999; 99US-0137694P.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Bukh J, Yanagi M, Emerson SU, Purcell RH;

XX WPI; 2001-091214/10.

XX New infectious nucleic acids of the GB virus-B clone, useful for indirectly studying the molecular properties of hepatitis C virus (HCV) and in developing vaccines and therapeutics for HCV.

XX Claim 3; Page 60-63; 96pp; English.

XX The present invention relates to GB virus-B. The nucleic acid molecules of the invention are useful for indirectly studying the molecular properties of hepatitis C virus (HCV). The infectious nucleic acid sequence of the GB virus-B clone and the HCV/GBV-B chimera may be used in the development of vaccines and therapeutics for HCV. (Updated on 11-SEP-2003 to standardise OS field)

```
SQ Sequence 9399 BP; 2125 A; 2326 C; 2430 G; 2518 T; 0 U; 0 Other;

Query Match      100.0%; Score 260; DB 4; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1e-77;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9140 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9199

QY 61 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 120
DB 9200 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 9259

QY 121 TGGGAAGCAGTCAGTATTAATCCCGTCGTGTGTGTGAGCGCTCAGCAGTATTTGTCCG 180
DB 9260 TGGGAAGCAGTCAGTATTAATCCCGTCGTGTGTGTGAGCGCTCAGCAGTATTTGTCCG 9319

QY 181 CTGTGCAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTGTCCCAAGCGGAGGCAACC 240
DB 9320 CTGTGCAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTGTCCCAAGCGGAGGCAACC 9379

QY 241 CCCGCTTGGAAATTAATAACT 260
DB 9380 CCCGCTTGGAAATTAATAACT 9399

RESULT 2
ID ADJ56732 standard; cDNA; 9399 BP.
AC ADJ56732;
XX
XX
XX 06-MAY-2004 (first entry)
XX
XX Genome length hepatitis GB virus B cDNA SeqID 2.
XX
XX ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
XX vaccine; virucidal; antiinflammatory.
XX
XX Hepatitis GB virus B.
XX
XX Key Location/Qualifiers
FH stem_loop 4..21
FT /*tag= a
FT stem_loop 29..61
FT /*tag= b
FT misc_binding 63..73
FT /*tag= c
FT /*bound_moiety= "GBV-B DNA bases 236-226"
FT stem_loop 81..131
FT /*tag= d
FT stem_loop 135..154
FT /*tag= e
FT stem_loop 156..218
FT /*tag= f
FT misc_binding 226..236
FT /*tag= g
FT /*bound_moiety= "GBV-B DNA bases 73-63"
FT stem_binding 242..250
FT /*tag= h
FT /*bound_moiety= "GBV-B DNA bases 427-419"
FT misc_binding 251..257
FT /*tag= i
FT /*bound_moiety= "GBV-B DNA bases 394-388"
FT misc_binding 258..270
FT /*tag= j
FT /*bound_moiety= "GBV-B DNA bases 362-349"
FT stem_loop 273..287
FT /*tag= k
FT stem_loop 288..337
FT /*tag= l
FT stem_loop 339..348
FT ..
```

```
FT misc_binding /*tag= m
FT 349..362
FT /*tag= n
FT /*bound_moiety= "GBV-B DNA bases 270-258"
FT stem_loop 363..384
FT /*tag= o
FT misc_binding 388..394
FT /*tag= p
FT /*bound_moiety= "GBV-B DNA bases 257-251"
FT stem_loop 395..406
FT /*tag= q
FT misc_binding 410..415
FT /*tag= r
FT misc_binding 419..427
FT /*tag= s
FT /*bound_moiety= "GBV-B DNA bases 434-429"
FT stem_loop 429..434
FT /*tag= t
FT /*bound_moiety= "GBV-B DNA bases 415-410"
FT stem_loop 435..459
FT /*tag= u
XX
XX WO2004005498-A1.
XX
XX 15-JAN-2004.
XX
XX 02-JUL-2003; 2003WO-US021002.
XX
XX 03-JUL-2002; 2002US-00189359.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX (INSP ) INST PASTEUR.
XX
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX WPI; 2004-091362/09.
XX
XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
XX polynucleotides. Specifically, it refers to using the hepatotropic
XX flavivirus GBV-B that has a unique phylogenetic relationship to the human
XX hepatitis C virus (HCV) and can serve as a surrogate virus in drug
XX discovery efforts related to antiviral drug development. The present
XX invention describes the construction of an infectious molecular clone
XX using the newly determined 3' terminal sequence of GBV-B. Furthermore,
XX the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
XX envelope proteins such that they can have utility as a vaccine immunogen
XX for hepatitis C. In addition, they can be used for screening compounds
XX active against viral infection, as well as for developing HCV
XX preventative and therapeutic treatments. Accordingly, these compositions
XX exhibit virucidal, antiinflammatory and hepatotropic activities. This
XX polynucleotide sequence is the GBV-B cDNA sequence of the invention.
XX
XX Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;
XX

Query Match      100.0%; Score 260; DB 12; Length 9399;
Best Local Similarity 100.0%; Pred. No. 1e-77;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9140 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9199

QY 61 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 120
DB 9200 GACGTCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGGCTGTG 9259
```

QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAGCGCTCAGACGTATATTTGTCGG 180
 DB 9260 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAGCGCTCAGACGTATATTTGTCGG 9319
 QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
 DB 9320 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
 QY 241 CCGCGTTGGAATTAATAAACT 260
 DB 9380 CCGCGTTGGAATTAATAAACT 9399

RESULT 3
 ADJ64244
 ID ADJ64244 standard; DNA; 9399 BP.
 XX AC ADJ64244;
 XX DT 20-MAY-2004 (first entry)
 XX GB virus B 3' terminal polynucleotide fragment seqid 2.
 XX antiinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
 KW hepatitis C virus; HCV; 3' terminal; ds.
 XX Hepatitis GB virus B.
 XX US2004039187-A1.
 XX 26-FEB-2004.
 XX 03-JUL-2002; 2002US-00189359.
 XX 04-JUN-1999; 99US-0137665P.
 XX 05-JUN-2000; 2000US-00587653.
 XX (TEXA) UNIV TEXAS SYSTEM.
 PA (INSP) INST PASTEUR.
 XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
 PI WPI; 2004-203294/19.
 XX New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
 PT diagnosing and in treating HCV and in investigating the mechanisms for
 PT the different biological properties of the viruses.

PS Claim 10; SEQ ID NO 2; 58pp; English.
 XX The invention describes a new isolated polynucleotide (I) encoding a 3'
 CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
 CC GBV-B genome, where at least part, but not all of a 5' nontranslated
 CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
 CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
 CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
 CC or chimeras are useful diagnosing or treating hepatitis C virus (HCV)
 CC and in investigating the mechanisms for the different biological
 CC properties of the viruses. This sequence represents a Hepatitis GB virus
 CC B (GBV-B) 3' terminal polynucleotide.
 XX Sequence 9399 BP; 2124 A; 2322 C; 2432 G; 2521 T; 0 U; 0 Other;

Query Match 100.0%; Score 260; DB 12; Length 9399;
 Best Local Similarity 100.0%; Pred. No. 1e-77;
 Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGTTTGGGACCATGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 60
 DB 9140 GAGTTTGGGACCATGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 9199
 QY 61 GAGCTCCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGAGTCTTCGGCTGTG 120

DB 9200 GAGCTCCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGAGTCTTCGGCTGTG 9259
 QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAGCGCTCAGACGTATATTTGTCGG 180
 DB 9260 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAGCGCTCAGACGTATATTTGTCGG 9319
 QY 181 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 240
 DB 9320 CTGTGCAGAGCGTAGTACCAAGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACC 9379
 QY 241 CCGCGTTGGAATTAATAAACT 260
 DB 9380 CCGCGTTGGAATTAATAAACT 9399

RESULT 4
 AAC92010
 ID AAC92010 standard; DNA; 309 BP.
 XX AC AAC92010;
 XX DT 20-MAR-2001 (first entry)
 XX GBV-B 3'X sequence and 3'UTR region.
 DE GBV-B; hepatitis GB viral-B; hepatitis C virus activity; ds.
 KW GBV-B; hepatitis GB viral-B; hepatitis C virus activity; ds.
 XX Hepatitis virus.
 XX WO200073466-A1.
 XX 07-DEC-2000.
 XX 22-MAY-2000; 2000WO-EP004622.
 XX 27-MAY-1999; 99GB-00012432.
 XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
 PA Traboni C;
 PI WPI; 2001-061544/07.
 XX Novel GB viral sequence useful in in vivo assays to identify agents which
 PT modulate hepatitis C virus activity.
 XX Example 2; Fig 3; 76pp; English.

CC The present invention relates to a hepatitis GB viral-B (GBV-B)
 CC polynucleotide (see AAC91998 and AAC91996). The hepatitis GB agent was
 CC first discovered by the inoculation of tamarins with serum from a
 CC patient, whose initials were GB, affected by acute hepatitis. The serum
 CC induced hepatitis in all inoculated tamarins. The present sequence
 CC comprises the sequence in AAC91998 (the 3' terminus of GBV-B, plus the 3'
 CC UTR. This sequence confers infectivity in tamarins on otherwise non-
 CC infective GBV-B genome. This sequence is useful in in vivo assays to
 CC identify agents which modulate hepatitis C virus (HCV) activity
 XX Sequence 309 BP; 64 A; 80 C; 100 G; 65 T; 0 U; 0 Other;

Query Match 99.4%; Score 258.4; DB 4; Length 309;
 Best Local Similarity 99.6%; Pred. No. 1.1e-77;
 Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GAGTTTGGGACCATGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 60
 DB 50 GAGTTTGGGACCATGTGATCAGAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 109
 QY 61 GAGCTCCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGAGTCTTCGGCTGTG 120
 DB 110 GAGCTCCCTCTTCGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGAGTCTTCGGCTGTG 169
 QY 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAGCGCTCAGACGTATATTTGTCGG 180

Db 170 TGGGAAGCAGTCAGTATTAATCCCGCTCGTGTGTGGTACGCTCAAGACGTACTTGTCCG 229

Qy 181 CTGTGCAGAGCGTACTTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGGCAACC 240

Db 230 CTGTGCAGAGCGTAGTACCAAGGGTGCACCCCGGTTTTTGTTCGAAGCGGAGGGCAACC 289

Qy 241 CCCGCTTGGAAATTAAAACT 260

Db 290 CCCGCTTGGAAATTAAAACT 309

RESULT 5
ADA77751
ID ADA77751 standard; RNA; 8069 BP.
XX
XX ADA77751;
XX AC
XX
DT 20-NOV-2003 (first entry)

DE Hepatitis GB virus B subgenomic neo-RepD replicon RNA sequence.
XX
KW neo-RepD; GB virus-B; GSV-B; replicon; autonomous replication; NS3-NS5B;
KW HCV; hepatitis C virus; antiviral; anti-HCV;
KW neomycin phosphotransferase gene; neo; ss.

OS Hepatitis GB virus B.

Key	Location/Qualifiers
5'UTR	1..445
FT	/*tag= a
FT	446..1315
FT	/*tag= b
FT	/product= "Core-neo fusion protein"
FT	1324..1934
FT	misc signal

FI	/tag= C	
FT	/label= IRES	
FT	/note= "IRES= Internal ribosome entry site of the	
FT	encephalomyocarditis virus, drives translation of the GSV	
FT	-B NS region"	
FT	1935..7712	CDS
FT		

```

F1      /-tag= d
FT      /product= "GBV-B polypeptide NS3-NS5B"
FT      /note= "NS= non structural protein"
FT      /

```

FT	CDS	5940..7712	
FT		/*tag= e	
FT		/partial	
FT		/product= "Non structural protein 5B (NS5B)"	
FT		/note= "NS5B is an RNA dependent RNA polymerase; Start	
FT		codon is absent"	
FT	3'UTR	7710..8069	
FT		/*tag= f	

XX
ND
CM - 170830C000C0M

XX
PP
24-111-2003

XX
DE 12-JAN-2003. 2003WC-EP000201

XX
DB 1E-JAN-2003 - 200308-03485730

PR 06-JUN-2002; 2002US-0386655P.
YY

PA (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
VY

PI De Tomassi A, Graziani R, Paonessa G, Traboni C;
vv

DR WPI; 2003-598503/56.

New GB virus B (GBV-B) replicon for identifying compounds that inhibit GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence, and a GBV-B 3' UTR.

PS Claim 1; Fig 1; 81pp; English.

CC This invention relates to a novel GB virus-B (GBV-B) replicon and
CC replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
CC autonomously replicate in a cultured cell to produce detectable levels of
CC one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5',
CC UTR, GBV-B structural region, selection or reporter sequence, internal
CC ribosome entry site, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
CC they are useful in providing tools for studying GBV-B replication,
CC polypeptide production and processing, identifying compounds that inhibit
CC GBV-B, providing a surrogate model for identifying compounds that inhibit
CC HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
CC Due to the similarity between GBV-B and the hepatitis C virus (HCV),
CC compounds that inhibit GBV-B may be useful antiviral agents, specifically
CC anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
CC neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
CC replacing the regions coding for structural proteins and the NS2 protein
CC with the sequences of mycristin phosphotransferase gene (neo) and
CC encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
CC the plasmid FL3/pACYC177. This polynucleotide sequence is the subgenomic
CC neo-RepD replicon sequence of the invention.

SQ Sequence 8069 BP; 1863 A; 2075 C; 2136 G; 0 T; 1995 U; 0 Other;
 Query Match 99.4%; Score 258.4; DB 9; Length 8069;
 Best Local Similarity 76.2%; Pred. No. 3.4e-77;
 Matches 199; Conservative 61; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTCGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCCATGGTCTTGAAAGGGAT 60

DH GAGTTCGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCCATGGTCTTGAAAGGGAT 7869

[illegible]

QY 121 TGGGAACAGTCAGTATAATCCCGTCGTGCTGGTCAGCGCTCACGACGTAATTGTCCG 180

QY 181 CTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGGCAACC 240

QY	241	CCGCGTTGGAA	TAAAAACT	260
1990			:::	

DD 8050 CCCGCCUUGGAAUUAAAAACU 8069

DECUIT F 5

ADA77752
ID ADA77752 standard; cDNA; 9397 BP.
...

AC ADA77752;

DT 20-NOV-2003 (first entry)

DE Genomic hepatitis GB virus B (GBV-B) replicon cDNA sequence.

KW GB virus-B; GBV-B; replicon; autonomous replication; NS3-NS5B; HCV;
KW hepatitis C virus; antiviral; anti-HCV; neomycin phosphotransferase gene;
KW neo; ss.

OS Hepatitis GB virus B.

Key	Location/Qualifiers
5'UTR	1..445
ET	/*tag= a
ET	445..9040
ET	/*tag= b
ET	/product= "GBV-B polyprotein"
ET	/note= "From core protein to non structural protein 5B"
ET	7268..9040
ET	CDS

```
FT FT /*tag= C
FT FT /partial
FT FT /product= "Non structural protein 5B (NS5B)"
FT FT /note= "NS5B is an RNA dependent RNA polymerase; start
FT FT codon is absent"
FT FT 9038..9397
FT FT /*tag= d
XX XX WO2003059944-A2.
XX XX 24-JUL-2003.
XX XX 13-JAN-2003; 2003WO-EP000281.
XX XX 15-JAN-2002; 2002US-0348573P.
XX XX 06-JUN-2002; 2002US-0386655P.
XX XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX XX De Tomasi A, Graziani R, Paonessa G, Traboni C;
XX XX WPI; 2003-598503/56.
XX XX P-PSDB; ADA77730.
XX XX New GB virus B (GBV-B) replicon for identifying compounds that inhibit
XX XX GBV-B or hepatitis C virus, comprises a GBV-B 5' UTR, a selection or
XX XX reporter sequence, an internal ribosome entry site, an NS3-NS5B sequence,
XX XX and a GBV-B 3' UTR.
XX XX Claim 3; Fig 2; 81pp; English.
XX XX This invention relates to a novel GB virus-B (GBV-B) replicon and
XX XX replicon enhanced cells. A GBV-B replicon is an RNA molecule able to
XX XX autonomously replicate in a cultured cell to produce detectable levels of
XX XX one or more GBV-B proteins. Specifically, it may comprise the GBV-B 5'
XX XX UTR, GBV-B structural region, selection or reporter sequence, internal
XX XX ribosome entry site, NS3-NS5B sequence, and GBV-B 3' UTR. Accordingly,
XX XX they are useful in providing tools for studying GBV-B replication,
XX XX polypeptide production and processing, identifying compounds that inhibit
XX XX GBV-B, providing a surrogate model for identifying compounds that inhibit
XX XX HCV, and providing a scaffold for producing GBV-B/HCV chimeric replicons.
XX XX Due to the similarity between GBV-B and the hepatitis C virus (HCV),
XX XX compounds that inhibit GBV-B may be useful antiviral agents, specifically
XX XX anti-HCV agents. The GBV-B subgenomic replicon constructs termed GBV-B-
XX XX neo-RepA (neo-RepA), neo-RepB, neo-RepC and neo-RepD were produced by
XX XX replacing the regions coding for structural proteins and the NS2 protein
XX XX with the sequences of neomycin phosphotransferase gene (neo) and
XX XX encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) in
XX XX the plasmid FL3/pACYC177. This polynucleotide sequence is the genomic GBV
XX XX -B replicon cDNA sequence of the invention.
XX XX Sequence 9397 BP; 2125 A; 2331 C; 2431 G; 2510 T; 0 U; 0 Other;
SQ

Query Match 99.4%; Score 258.4; DB 9; Length 9397;
Best Local Similarity 99.6%; Pred. No. 3.6e-77;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCACCATTGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9138 GAGTTTGGCACCATTGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 9197

QY 61 GAGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTGGCTGTG 120
DB 9198 GAGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTGGCTGTG 9257

QY 121 TGGGAACAGTCAGTATAATTCCTCGTGTGTGGTGAACGCTCAGCGCTCAGACGATATTGTCCG 180
DB 9258 TGGGAACAGTCAGTATAATTCCTCGTGTGTGGTGAACGCTCAGCGCTCAGACGATATTGTCCG 9317

QY 181 CTGTGCAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 240
DB 9318 CTGTGCAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 9377
```

```
QY 241 CCGGCTTGGAAATTAATAAACT 260
DB 9378 CCGGCTTGGAAATTAATAAACT 9397

RESULT 7
AAC91998
ID AAC91998 standard; DNA; 259 BP.
XX AAC91998;
XX 20-MAR-2001 (first entry)
XX GBV-B 3'X DNA sequence.
XX GBV-B; hepatitis GB viral-B; hepatitis C virus activity; da.
XX Hepatitis virus.
XX WO200073466-A1.
XX 07-DEC-2000.
XX 22-MAY-2000; 2000WO-EP004622.
XX 27-MAY-1999; 99GB-00012432.
XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.
XX Traboni C;
XX WPI; 2001-061544/07.
XX Novel GB viral sequence useful in in vivo assays to identify agents which
XX modulate hepatitis C virus activity.
XX Claim 1; Page 14; 76pp; English.
XX The present sequence is an isolated hepatitis GB viral-B (GBV-B)
XX polynucleotide. The hepatitis GB agent was first discovered by the
XX inoculation of tamarins with serum from a patient, whose initials were
XX GB, affected by acute hepatitis. The serum induced hepatitis in all
XX inoculated tamarins. The present sequence comprises the 3' terminus of
XX GBV-B, and confers infectivity in tamarins on otherwise non-infective GBV
XX -B genome. This sequence is useful in in vivo assays to identify agents
XX which modulate hepatitis C virus (HCV) activity
XX Sequence 259 BP; 53 A; 64 C; 81 G; 61 T; 0 U; 0 Other;
SQ

Query Match 99.0%; Score 257.4; DB 4; Length 259;
Best Local Similarity 99.6%; Pred. No. 2.2e-77;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCACCATTGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
DB 1 AGTTTGGCACCATTGTGTGATCAGAACCCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60

QY 62 ACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTGT 121
DB 61 ACGTCCCTTCTGGCTCATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGCTGTGT 120

QY 122 GGGGAACAGTCAGTATAATTCCTCGTGTGTGGTGAACGCTCAGCGCTCAGACGATATTGTCCGC 181
DB 121 GGGGAACAGTCAGTATAATTCCTCGTGTGTGGTGAACGCTCAGCGCTCAGACGATATTGTCCGC 180

QY 182 TGTCGAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 241
DB 181 TGTCGAGAGCGTAGTACCAAGGCGTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 240

QY 242 CCGCTTGGAAATTAATAAACT 260
DB 241 CCGCTTGGAAATTAATAAACT 259
```

```
RESULT 8
ADJ56731
ID ADJ56731 standard; RNA; 260 BP.
XX AC
XX ADJ56731;
XX DT
XX 06-MAY-2004 (first entry)
XX DE
XX 3' terminal sequence of the hepatitis GB virus B RNA SeqID 1.
XX ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
XX vaccine; virucidal; antiinflammatory.
XX OS
XX Hepatitis GB virus B.
XX FH Key
XX Location/Qualifiers
XX misc_binding
XX 152..156
XX /tag= a
XX /bound_moiety= "3' terminal GBV-B RNA bases 199-195"
XX misc_binding
XX 157..160
XX /tag= b
XX /bound_moiety= "3' terminal GBV-B RNA bases 192-189"
XX misc_binding
XX 161..162
XX /tag= c
XX /bound_moiety= "3' terminal GBV-B RNA bases 187-186"
XX misc_binding
XX 163..166
XX /tag= d
XX /bound_moiety= "3' terminal GBV-B RNA bases 184-181"
XX stem_loop
XX 168..179
XX /tag= e
XX /label= SL-3
XX misc_binding
XX 181..184
XX /tag= f
XX /bound_moiety= "3' terminal GBV-B RNA bases 166-163"
XX misc_binding
XX 186..187
XX /tag= g
XX /bound_moiety= "3' terminal GBV-B RNA bases 162-161"
XX misc_binding
XX 189..192
XX /tag= h
XX /bound_moiety= "3' terminal GBV-B RNA bases 160-157"
XX misc_binding
XX 195..199
XX /tag= i
XX /bound_moiety= "3' terminal GBV-B RNA bases 156-152"
XX stem_loop
XX 201..212
XX /tag= j
XX /label= SL-2
XX 214..220
XX /tag= k
XX /bound_moiety= "3' terminal GBV-B RNA bases 258-252"
XX misc_binding
XX 221..230
XX /tag= l
XX /bound_moiety= "3' terminal GBV-B RNA bases 250-241"
XX stem_loop
XX 232..239
XX /tag= m
XX /label= SL-1
XX misc_binding
XX 241..250
XX /tag= n
XX /bound_moiety= "3' terminal GBV-B RNA bases 230-221"
XX misc_binding
XX 252..258
XX /tag= o
XX /bound_moiety= "3' terminal GBV-B RNA bases 220-214"
XX WO2004005498-A1.
XX 15-JAN-2004.
XX 02-JUL-2003; 2003WO-US021002.
XX 03-JUL-2002; 2002US-00189359.
XX (TEXA ) UNIV TEXAS SYSTEM.
XX (INSP ) INST PASTEUR.
XX PI
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX DR
XX WPI; 2004-091362/09.
XX PT
XX New chimeric GBV-B polynucleotide, useful as a model for hepatitis C
XX virus, for identifying compounds active against a viral infection, or for
XX developing hepatitis C virus preventive and therapeutic treatments.
XX PS
XX Example 1; SEQ ID NO 1; 108pp; English.
XX CC
XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
XX polynucleotides. Specifically, it refers to using the hepatotropic
XX flavivirus GBV-B that has a unique phylogenetic relationship to the human
XX hepatitis C virus (HCV) and can serve as a surrogate virus in drug
XX discovery efforts related to antiviral drug development. The present
XX invention describes the construction of an infectious molecular clone
XX using the newly determined 3' terminal sequence of GBV-B. Furthermore,
XX the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
XX envelope proteins such that they can have utility as a vaccine immunogen
XX for hepatitis C. In addition, they can be used for screening HCV
XX active against viral infection, as well as for developing HCV
XX preventative and therapeutic treatments. Accordingly, these compositions
XX exhibit virucidal, antiinflammatory and hepatotropic activities. This
XX polynucleotide sequence is the 3' terminal sequence of GBV-B RNA sequence
XX of the invention.
XX SQ
XX Sequence 260 BP; 53 A; 63 C; 82 G; 0 T; 62 U; 0 Other;
XX
XX Query Match 97.8%; Score 254.4; DB 12; Length 260;
XX Best Local Similarity 75.8%; Pred. No. 2.3e-76;
XX Matches 194; Conservative 61; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 5 TTGGCGACATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTCAAGGGGATGACG 64
XX ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX Db 4 UUUGCGACCAUGGUGGUAUCAGAACCGUUUCGGUGAAGCCAUUGGUCUGAAGGGGAUGCG 63
XX
XX QY 65 TCCCTTCCTGGCTCATCCACAAAACCGTCTCGGGTGGTGAGGAGTCTCGCTGTGTGGG 124
XX ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX Db 64 UCCCCUUGGUGGUAUUAUCCGACAAAACCGUCUGGUGGUGAGAGUCCUGGUCUGUGGG 123
XX
XX QY 125 AAGCAGTCAGTATATTCCTCGTGTGTGGTGGAGCCCTCAGACGTATTTGTCGCTGT 184
XX ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX Db 124 AAGCAGUCAGUAUAUUAUCCGUGUGUGAGCGCCUACAGAGUAUUUGUCCGUGU 183
XX
XX QY 185 GCAGAGCGTAGTACCAAGGGCTGCACCCCGTTTGTTCAGCGGAGGGCAACCCCG 244
XX ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX Db 184 GCAGAGCGUAGUACCAAGGGCUGCACCCCGUUUUUUGUCCAAAGCGGAGGCAACCCCG 243
XX
XX QY 245 CTTGGAATTAAAAACT 260
XX ::::::::::::::::::::::
XX Db 244 CUUGGAUUUAAAAACU 259
XX
XX RESULT 9
XX ADJ64243
XX ID ADJ64243 standard; DNA; 260 BP.
XX AC
XX ADJ64243;
XX DT
XX 20-MAY-2004 (first entry)
XX DE
XX GB virus B 3' terminal polynucleotide fragment seqid 1.
XX antinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
XX hepatitis C virus; HCV; 3' terminal; ds.
XX OS
XX Hepatitis GB virus B.
XX XX
XX US2004039187-A1.
XX PD
XX 26-FEB-2004.
XX XX
```


XX The invention describes a new isolated polynucleotide (I) encoding a 3'
CC sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
CC GBV-B genome, where at least part, but not all of a 5' nontranslated
CC region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
CC (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
CC fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
CC or chimaeras are useful diagnosing or treating hepatitis C virus (HCV)
CC and in investigating the mechanisms for the different biological
CC properties of the viruses. This sequence represents a Hepatitis GB virus
CC B (GBV-B) 3' terminal polynucleotide. Note: This sequence represents
XX nucleotides 152-259 of the sequence shown in ADJ64243.

SQ Sequence 108 BP; 23 A; 30 C; 30 G; 0 T; 25 U; 0 Other;

Query Match 41.5%; Score 108; DB 12; Length 108;
Best Local Similarity 76.9%; Pred. No. 1.7e-26;
Matches 83; Conservative 25; Mismatches 0; Indels 0; Gaps 0;

QY 153 TGGTGACGCTCACGACGTATTGTCCGCTGTGCAGCGCTAGTACCAAGGGCTGCACCC 212
DB 1 UGGUGAGCCUCACGACGUUUUUGCGUGGAGCGAGCGUAGUACCAAGGGCGCACCC 60

QY 213 CGGTTTTTGTCCACGCGAGGCGCAACCCCGCTCGAATTAATAACT 260
DB 61 CGGUUUUGUCCAGCGAGGCGCAACCCCGCGUUGGAUUUAAAAACU 108

RESULT 13

AAC92012
ID AAC92012 standard; RNA; 82 BP.

XX AAC92012;

DT 20-MAR-2001 (first entry)

DE GBV-B 3'X RNA sequence #2.

XX GBV-B; hepatitis GB viral-B; ds.

XX Hepatitis virus.

XX WO200073466-A1.

XX 07-DEC-2000.

XX 22-MAY-2000; 2000WO-EP004622.

XX 27-MAY-1999; 99GB-00012432.

XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX Traboni C;

XX WPI; 2001-061544/07.

XX Novel GB viral sequence useful in in vivo assays to identify agents which
PT modulate hepatitis C virus activity.

XX Disclosure; Fig 4; 76pp; English.

XX The present sequence is an isolated hepatitis GB viral-B (GBV-B)
CC polynucleotide. The hepatitis GB agent was first discovered by the
CC inoculation of tamarins with serum from a patient, whose initials were
CC GB, affected by acute hepatitis. The serum induced hepatitis in all
CC inoculated tamarins. The present sequence comprises the 3' terminus of
CC GBV-B

XX Sequence 82 BP; 19 A; 23 C; 23 G; 0 T; 17 U; 0 Other;

Query Match 31.5%; Score 82; DB 4; Length 82;
Best Local Similarity 79.3%; Pred. No. 1.1e-17;
Matches 65; Conservative 17; Mismatches 0; Indels 0; Gaps 0;

QY 179 CGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTGTTCGAAGCGAGGGCAA 238
DB 1 CGCGUGGAGAGCGUAGUACCAAGGGCGUGCACCCCGUUUUUUGUCCAAAGCGAGGGCAA 60

QY 239 CCCCGCGCTTGGATTAATAACT 260
DB 61 CCCCGCGUUGGAUUUAAAAACU 82

RESULT 14

ADJ56742/c
ID ADJ56742 standard; DNA; 108 BP.

XX ADJ56742;

XX 06-MAY-2004 (first entry)

XX Mutagenic PCR primer used to amplify hepatitis GB virus B cDNA SeqID 12.

XX ss; GB virus B; GBV-B; HCV; flavivirus; hepatitis C virus; antiviral;
KW vaccine; virucidal; antiinflammatory; PCR; primer; mutagenic.

XX Hepatitis GB virus B.

OS Synthetic.

XX WO2004005498-A1.

XX 15-JAN-2004.

XX 02-JUL-2003; 2003WO-US021002.

XX 03-JUL-2002; 2002US-00189359.

XX (TEXA) UNIV TEXAS SYSTEM.

XX (INSP) INST PASTEUR.

XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;

XX WPI; 2004-091362/09.

XX New chimeric GBV-B polynucleotide, useful as a model for hepatitis C
PT virus, for identifying compounds active against a viral infection, or for
PT developing hepatitis C virus preventive and therapeutic treatments.

XX Example 22; SEQ ID NO 12; 108pp; English.

XX This invention relates to novel isolated chimeric GB virus B (GBV-B)/HCV
CC polynucleotides. Specifically, it refers to using the hepatotropic
CC flavivirus GBV-B that has a unique phylogenetic relationship to the human
CC hepatitis C virus (HCV) and can serve as a surrogate virus in drug
CC discovery efforts related to antiviral drug development. The present
CC invention describes the construction of an infectious molecular clone
CC using the newly determined 3' terminal sequence of GBV-B. Furthermore,
CC the GBV-B/HCV chimeras exhibit liver-specific expression and express HCV
CC envelope proteins such that they can have utility as a vaccine immunogen
CC for hepatitis C. In addition, they can be used for screening compounds
CC active against viral infection, as well as for developing HCV
CC preventative and therapeutic treatments. Accordingly, these compositions
CC exhibit virucidal, antiinflammatory and hepatotropic activities. This
CC oligonucleotide sequence is a mutagenic PCR primer used to amplify the
CC GBV-B cDNA sequence of the invention.

XX Sequence 108 BP; 23 A; 33 C; 30 G; 22 T; 0 U; 0 Other;

Query Match 25.4%; Score 66; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.7e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GTACCAAGGCTGCACCCCGTTTTTGTTCGAAGCGAGGGCAACCCCGCTTGGAAATTA 254

DB 108 GTACCAAGGCTGCACCCCGTTTTTGTTCGAAGCGAGGGCAACCCCGCTTGGAAATTA 49

```

QY 255 AAAACT 260
Db 48 AAAACT 43

RESULT 15
ADJ64254/c
ID ADJ64254 standard; DNA; 108 BP.
XX
AC ADJ64254;
XX
XX 20-MAY-2004 (first entry)
XX
DE Hepatitis GB virus B infectious clone related primer seqid 12.
XX
XX antinflammatory; hepatotropic; virucide; GB virus B; GBV-B;
KW hepatitis C virus; HCV; 3' terminal; primer; ss; PCR; infectious clone.
XX
XX Hepatitis GB virus B.
XX
XX US2004039187-A1.
XX
XX 26-FEB-2004.
XX
XX 03-JUL-2002; 2002US-00189359.
XX
XX 04-JUN-1999; 99US-0137665P.
XX
XX 05-JUN-2000; 2000US-00587653.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX (INSP ) INST PASTEUR.
XX
XX Martin A, Sangar DV, Lemon SM, Rijnbrand R;
XX
XX WPI; 2004-203294/19.
XX
XX New GB virus B and/or hepatitis C virus (HCV) sequences, useful in
XX diagnosing and in treating HCV and in investigating the mechanisms for
XX the different biological properties of the viruses.
XX
XX Example 22; SEQ ID NO 12; 58pp; English.
XX
XX The invention describes a new isolated polynucleotide (I) encoding a 3'
XX sequence of the GB virus B (GBV-B) genome, or which comprises a chimeric
XX GBV-B genome, where at least part, but not all of a 5' nontranslated
XX region (NTR) sequence is derived from a hepatitis C virus (HCV) 5' NTR.
XX (I) is a GB virus B and/or hepatitis C virus polynucleotide comprising a
XX fully defined of 260 or 9399 bp (SEQ ID NOS: 1 or 2). The polynucleotides
XX or chimaeras are useful diagnosing or treating hepatitis C virus (HCV)
XX and in investigating the mechanisms for the different biological
XX properties of the viruses. This sequence represents a primer used in the
XX construction of a Hepatitis GB virus B (GBV-B) infectious clone.
XX
XX Sequence 108 BP; 23 A; 33 C; 30 G; 22 T; 0 U; 0 Other;

Query Match 25.4%; Score 66; DB 12; Length 108;
Best Local Similarity 100.0%; Pred. No. 3.7e-12;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GTACCAAGGGCTGCACCCGGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 254
Db 108 GTACCAAGGGCTGCACCCGGTTTGTTCGAAGCGGAGGCAACCCCGCTTGGAAATTA 49

QY 255 AAAACT 260
Db 48 AAAACT 43

```

Search completed: December 7, 2005, 21:08:27
Job time : 370 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 13:50:44 ; Search time 127 Seconds
(without alignments)
3639.103 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gaggttggcgaccatggtgg.....cccgcttgaataaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/1 COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5 COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/H COMB.seq.*
6: /cgn2_6/ptodata/1/ina/PCBUS COMB.seq.*
7: /cgn2_6/ptodata/1/ina/PP COMB.seq.*
8: /cgn2_6/ptodata/1/ina/RE COMB.seq.*
9: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	258.4	99.4	309	US-09-579-302-1	Sequence 1, Appli
2	257.4	99.0	259	US-09-579-302-2	Sequence 2, Appli
3	257.4	99.0	259	US-09-579-302-3	Sequence 3, Appli
4	257.4	99.0	259	US-09-579-302-4	Sequence 4, Appli
5	257.4	99.0	259	US-09-579-302-5	Sequence 5, Appli
6	82	31.5	82	US-09-579-302-14	Sequence 14, Appl
7	30.8	11.8	133358	US-09-949-016-16964	Sequence 16964, A
8	30.8	11.8	133360	US-09-949-016-12651	Sequence 12651, A
9	30.4	11.7	20444	US-09-949-016-15750	Sequence 15750, A
10	29.8	11.5	601	US-09-949-016-182173	Sequence 182173,
11	29.8	11.5	24183	US-08-943-731-3	Sequence 3, Appli
12	29.4	11.3	1002	US-09-468-253B-53	Sequence 53, Appl
13	29.4	11.3	1005	US-09-468-253B-51	Sequence 51, Appl
14	29.4	11.3	1047	US-09-468-253B-61	Sequence 61, Appl
15	29.4	11.3	1047	US-09-468-253B-62	Sequence 62, Appl
16	29.4	11.3	2092	US-09-468-253B-13	Sequence 13, Appl
17	29.4	11.3	2473	US-09-468-253B-15	Sequence 15, Appl
18	29.2	11.2	1368	US-07-677-413-1	Sequence 1, Appli
19	29.2	11.2	1571	US-07-677-413-2	Sequence 2, Appli
20	29.2	11.2	1830121	US-09-557-884-1	Sequence 1, Appli
21	29.2	11.2	1830121	US-09-643-990A-1	Sequence 1, Appli
22	29.2	11.2	1830121	US-10-158-865-1	Sequence 1, Appli
23	29	11.2	1630	US-09-949-016-3210	Sequence 3210, Ap
24	29	11.2	1723	US-07-841-646-28	Sequence 28, Appl

25	29	11.2	1723	2	US-07-901-703-10	Sequence 10, Appl
26	29	11.2	1723	2	US-08-147-023-28	Sequence 28, Appl
27	29	11.2	1723	2	US-08-206-864-3	Sequence 3, Appli
28	29	11.2	1723	2	US-08-278-729A-20	Sequence 20, Appl
29	29	11.2	1723	2	US-08-480-528A-7	Sequence 7, Appli
30	29	11.2	1723	2	US-08-479-666-7	Sequence 7, Appli
31	29	11.2	1723	2	US-08-155-343A-20	Sequence 20, Appl
32	29	11.2	1723	2	US-08-408-672-20	Sequence 20, Appl
33	29	11.2	1723	2	US-08-643-563A-20	Sequence 20, Appl
34	29	11.2	1723	2	US-08-447-570-28	Sequence 20, Appl
35	29	11.2	1723	2	US-08-643-763A-20	Sequence 20, Appl
36	29	11.2	1723	2	US-08-462-623-20	Sequence 20, Appl
37	29	11.2	1723	2	US-08-451-953A-20	Sequence 20, Appl
38	29	11.2	1723	2	US-08-459-346-5	Sequence 5, Appli
39	29	11.2	1723	2	US-08-445-468A-20	Sequence 20, Appl
40	29	11.2	1723	2	US-08-901-200A-7	Sequence 7, Appli
41	29	11.2	1723	2	US-08-449-700-28	Sequence 28, Appl
42	29	11.2	1723	2	US-08-449-699A-28	Sequence 28, Appl
43	29	11.2	1723	2	US-08-461-397A-20	Sequence 20, Appl
44	29	11.2	1723	2	US-08-912-088-20	Sequence 20, Appl
45	29	11.2	1723	3	US-08-278-730A-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-09-579-302-1
; Sequence 1, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Newburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 309
; TYPE: DNA
; ORGANISM: GBV-B-like virus
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1)..(309)
; OTHER INFORMATION: cDNA complementary to the last 49 published 3'UTR
; OTHER INFORMATION: nucleotides of GBV-B plus a novel nucleotide (C)
; OTHER INFORMATION: in position 44 and 259 novel nucleotide sequence
; OTHER INFORMATION: at the 3' end.
US-09-579-302-1

Query Match 99.4%; Score 258.4; DB 3; Length 309;
Best Local Similarity 99.6%; Pred. No. 8.2e-81;
Matches 259; Conservative 1; Indels 0; Gaps 0;

Qy 1 GAGTTTGGCGACCATGTCGATCAGAAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 60
Db 50 GAGTTTGGCGACCATGTCGATCAGAAACCGTTTCGGGTGAAGCATGGTCTGAAGGGGAT 109
Qy 61 GAGTCCTCTTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTTGGCTGTG 120
Db 110 GAGTCCTCTTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTTGGCTGTG 169
Qy 121 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAAGCATGGTCTTGGCTGTG 180
Db 170 TGGGAAGCAGTCAGTATATAATTCCTCGTGTGTGGTGAAGCATGGTCTTGGCTGTG 229
Qy 181 CTGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTTCACAGCGAGGCAACC 240
Db 230 CTGTGAGAGCGTAGTACCAAGGGCTGCACCCCGTTTTTTTGTTCACAGCGAGGCAACC 289

```
QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 290 CCCGCTTGGAAATTAATAAACT 309

RESULT 2
US-09-579-302-2
; Sequence 2, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 259
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-2

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 76.1%; Pred. No. 1.7e-80;
Matches 197; Conservative 61; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
Db 1 AGUUGGCGACCAUGUGGAUCAGAACCGUUUCGGUGAAGCCAUUGUCUGAAGGGGAUG 60

QY 62 ACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTGT 121
Db 61 AGUUGCCUUCUGGCUCAUCCACAAAACCGUCUCGCGUGGUGAGGAGUCCUGGUGUGU 120

QY 122 GGGAGCAGTCAGTATTAATTCCTCGTGTGTGTGAGCGCTCAGCAGTATTTGTCCGC 181
Db 121 GGGAGCAGUCAGUAUAUUCGUGUGUGUGAGCGCCUCACGACGUAUCUUGUCCGC 180

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTCCAAAGCGGAGGCAACCC 241
Db 181 UUGUCAGAGCGUAGUACCAAGGGCGUCACCCCGUUUUUUUCCAAAGCGGAGGCAACCC 240

QY 242 CCGCTTGGAAATTAATAAACT 260
Db 241 CCGCUUGGAUUAATAAAACU 259

RESULT 3
US-09-579-302-3/c
; Sequence 3, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 259
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-3

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 290 CCCGCTTGGAAATTAATAAACT 309

RESULT 4
US-09-579-302-4
; Sequence 4, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 259
; TYPE: DNA
; ORGANISM: GBV-B-like virus
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (1)-(259)
; OTHER INFORMATION: cDNA of positive strand 3'UTR of GBV-B
US-09-579-302-4

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 61
Db 1 AGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATG 60

QY 62 ACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTGT 121
Db 61 ACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCCCTGGCTGTGT 120

QY 122 GGGAGCAGTCAGTATTAATTCCTCGTGTGTGTGAGCGCTCAGCAGTATTTGTCCGC 181
Db 121 GGGAGCAGTCAGTATTAATTCCTCGTGTGTGTGAGCGCTCAGCAGTATTTGTCCGC 180

QY 182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTCCAAAGCGGAGGCAACCC 241
Db 181 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTTCCAAAGCGGAGGCAACCC 240

QY 242 CCGCTTGGAAATTAATAAACT 260
Db 241 CCGCTTGGAAATTAATAAACT 259

RESULT 5
US-09-579-302-5/c
; Sequence 5, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 259
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-5/c

Query Match 99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```



```

; Sequence 5, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 259
; TYPE: DNA
; ORGANISM: GBV-B-like virus
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1)..(259)
; OTHER INFORMATION: cdna of negative strand 3'UTR of GBV-B
US-09-579-302-5

Query Match          99.0%; Score 257.4; DB 3; Length 259;
Best Local Similarity 99.6%; Pred. No. 1.7e-80;
Matches 258; Conservative 0; Mismatches 1; Indels 0; Gaps 0

Qy      2 AGTTTGGCGACCATGGTGGATCAGAACCGTTTTCGGGTGAAGCCCATGGTCTCAAGGGGATG 61
Db      259 AGTTTGGGACCATGGTGGATCAGAACCGTTTTCGGGTGAAGCCCATGGTCTCAAGGGGATG 200

Qy      62 ACGTCCCTTCTGGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGAGTCTCTGGGTGTGT 121
Db      199 ACGTCCCTTCTGGGCTCATCCACAAAAACCGTCTCGGTGGGTGAGAGTCTCTGGGTGTGT 140

Qy      122 GGGAGCAGTCAGTATATATCCCTCGTGTGTGTGAGCGCTCACGAGGTATTTGTCCGC 181
Db      139 GGGAGCAGTCAGTATATATCCCTCGTGTGTGTGAGCGCTCACGAGGTATTTGTCCGC 80

Qy      182 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 241
Db      79 TGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGCGGAGGCAACCC 20

Qy      242 CCGCTTGGAAATTAATAACT 260
Db      19 CCGCTTGGAAATTAATAACT 1

RESULT 6
US-09-579-302-14
; Sequence 14, Application US/09579302
; Patent No. 6627437
; GENERAL INFORMATION:
; APPLICANT: Traboni, Cinzia
; TITLE OF INVENTION: No. 6627437el GBV sequence
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/579,302
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: GB 9912432.3
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 82
; TYPE: RNA
; ORGANISM: GBV-B-like virus
US-09-579-302-14

Query Match          31.5%; Score 82; DB 3; Length 82;
Best Local Similarity 79.3%; Pred. No. 5.2e-19;
Matches 65; Conservative 17; Mismatches 0; Indels 0; Gaps 0

Qy      179 CGCTGTGCAGCGTAGTACCAAGGGCTGACCCCGTTTTTTGTTCGAAGCGGAGGCAAC 238

```



```
/
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-943-731-3

Query Match      11.5%; Score 29.8; DB 3; Length 24183;
Best Local Similarity 54.0%; Pred. No. 12;
Matches 61; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 121 TGGGAAGCAGTCAGTATATTCCTCGTCGTGTGGTGACGCTCAGACGTATTTGTCGCG 180
Db 4413 TTGACGTAGTCAGGGGACTCACCGTATTTCTGGCTGGCAGCTCATGGCAAGTTTCTCTC 4354

QY 181 CTGTGCAGACGCTAGTACCAAGGCTGCACCCCGTTTTTGTTCCTCAAGCGGAG 233
Db 4353 CTGGGCGCAGGGGCTCACATGGATCAGCATCCATTGAAGTTCAAACTGGAG 4301

RESULT 12
US-09-468-253B-53/c
; Sequence 53, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 53
; LENGTH: 1002
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-468-253B-53

Query Match      11.3%; Score 29.4; DB 3; Length 1002;
Best Local Similarity 50.2%; Pred. No. 4.5;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGTCCTCTTGGCTCATCCACAAAAACCGTCTC 95
Db 900 GGGGTATTCATAGGGTGTGTAGATTAAACCAGCTTCAGGCCCTGGAGGAGCTATTGCAGC 841

QY 96 GGGTGGGTGAGGAGTCTCGTGTGGGAAGCAGTCAGTATATTTCCCGTC-----GTG 150
Db 840 AGTTGGGTGAGGAGTTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAAAGGCAT 781

QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 780 TATGGTAGGGCCAGCTGGCGTAGGAGTACGAGGGCGAGCTGGTGGGAGAACCCGGCGCAGG 721

QY 211 CCGGGTTTTTGTTCCAAGCGGAG 233
Db 720 CCCAGTATGATCCTTTGGAGCAG 698

RESULT 13
US-09-468-253B-51/c
; Sequence 51, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-468-253B-61

Query Match      11.3%; Score 29.4; DB 3; Length 1047;
Best Local Similarity 50.2%; Pred. No. 4.6;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY 36 GGTGAAGCCATGCTCTGAAGGGGATGACGTCCTCTTGGCTCATCCACAAAAACCGTCTC 95
Db 878 GGGGTACTCATAGGGTGTATAGATTAAACCAGCTTCGGGCCCTGGAGGAACTATTTCAGC 819

QY 96 GGGTGGGTGAGGAGTCTCGTGTGGGAAGCAGTCAGTATATTTCCCGTC-----GTG 150
Db 818 AGTTGGGTGAGGAGTTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAAAGGCAT 759

QY 151 TGTGGTGACGCCCTCAGACGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCAC 210
Db 758 TATGGTAGGGCCAGCTGGCGTAGGAGTACGCGGGCAGCTGGTGGGAGAACCCGGCGCAGG 699

QY 211 CCGGGTTTTTGTTCCAAGCGGAG 233
Db 698 CCCAGTATGATCCTTTGGAGCAG 676

RESULT 15
US-09-468-253B-61/c
; Sequence 61, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-468-253B-61
```

```
US-09-468-253B-62/c
; Sequence 62, Application US/09468253B
; Patent No. 6750012
; GENERAL INFORMATION:
; APPLICANT: Gould-Rothberg, Bonnie
; TITLE OF INVENTION: Method of Identifying a Psychotropic Agent Using
; FILE OF INVENTION: Differential Gene Expression
; FILE REFERENCE: 15966-519
; CURRENT APPLICATION NUMBER: US/09/468,253B
; CURRENT FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: 60/113,127
; PRIOR FILING DATE: 1998-12-21
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 1047
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-468-253B-62

Query Match      11.3%; Score 29.4; DB 3; Length 1047;
Best Local Similarity 50.2%; Pred. No. 4.6;
Matches 102; Conservative 0; Mismatches 96; Indels 5; Gaps 1;

QY      36 GGTGAAGCCATGGTCTGAAGGGGATGACGTCCCTTCTGGCTCATCCACAAAACCGTCTC 95
Db      878 GGGTATTTCATAGGGTGTGTAGATTAAACCAGCTTCAGGCCCTGGAGGACTATTGCAGC 819

QY      96 GGGTCGGGTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATAATTCCCGTC-----GTG 150
Db      818 AGTTGGGTGAGGAGTTCCGTTTGGCATGACAGCGGTCTGTATTGTCTGATCAAAGGCAT 759

QY      151 TGTGTGACCGCTCAGGAGTATTGTCCGCTGTGAGAGCGTAGTACCAAGGCTGCAC 210
Db      758 TATGTAGGGCCAGCTGGCGTAGGTGTACGAGAGCAGCTGGTGGGAGGACAGGCGCAGG 699

QY      211 CCCGTTTTTTTGTTCGAAGCGGAG 233
Db      698 CCCAGTGATGATCCTTGGAGCAG 676
```

Search completed: December 7, 2005, 19:45:11
Job time : 130 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 17:19:59 ; Search time 535 Seconds
(without alignments)
4018.763 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 gsgtttggcagcatgtgg.....cccgcttgaataaaact 260

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 9793542 seqs, 413469005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:*
1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/prodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	260	100.0	9399	3	US-09-742-659-1 Sequence 1, Appli
2	260	100.0	9399	7	US-10-189-359-2 Sequence 2, Appli
3	258.4	99.4	8069	9	US-10-501-412-1 Sequence 1, Appli
4	258.4	99.4	9397	9	US-10-501-412-2 Sequence 2, Appli
5	254.4	97.8	260	7	US-10-189-359-1 Sequence 1, Appli
C 6	66	25.4	108	7	US-10-189-359-12 Sequence 12, Appli
C 7	32	12.3	3179	6	US-10-168-651-36 Sequence 36, Appli
8	31.8	12.2	645	5	US-10-027-632-132082 Sequence 132082,
9	31.8	12.2	645	6	US-10-027-632-132082 Sequence 132082,
10	31.4	12.1	645	4	US-09-925-065A-425966 Sequence 425966,
11	31	11.9	1095	6	US-10-000-897-36 Sequence 36, Appli
12	31	11.9	1095	8	US-10-818-168-36 Sequence 36, Appli
13	30.6	11.8	1150	4	US-09-925-065A-551953 Sequence 551953,
14	30.6	11.8	3772	10	US-11-097-143-31040 Sequence 31040, A
15	30.2	11.6	645	4	US-09-925-065A-425967 Sequence 425967,
16	30	11.5	381	8	US-10-425-115-30124 Sequence 30124, A
17	30	11.5	626	4	US-09-925-065A-487889 Sequence 487889,
18	30	11.5	2499	7	US-10-267-502-86 Sequence 86, Appli
19	30	11.5	3855	10	US-11-097-143-42398 Sequence 42398, A
20	30	11.5	4497	10	US-11-097-143-54233 Sequence 54233, Ap
C 21	30	11.5	10907	10	US-11-097-143-42397 Sequence 42397, A
C 22	30	11.5	13815	10	US-11-097-143-54222 Sequence 54222, Ap
C 23	29.8	11.5	600	9	US-10-972-079-75557 Sequence 75557, A

ALIGNMENTS

RESULT 1

US-09-742-659-1
; Sequence 1, Application US/09742659
; Patent No. US200010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 9399
; TYPE: DNA
; ORGANISM: GB virus-B
US-09-742-659-1

Query Match 100.0%; Score 260; DB 3; Length 9399;
Best Local Similarity 100.0%; Pred. No. 6.2e-82;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 9140 GAGTTTGGCGACCATGGTGGATCAGAACCGTTTCGGTGAAGCCATGGTCTGAAGGGGAT 9199
QY 61 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTTGGCTGTG 120
DB 9200 GAGCTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTTGGCTGTG 9259
QY 121 TGGGAAGCAGTCAGTATATAATCCCGTCGTGTGTGTGAGCGCTCAGACGATATTTGTCCG 180
DB 9260 TGGGAAGCAGTCAGTATATAATCCCGTCGTGTGTGTGAGCGCTCAGACGATATTTGTCCG 9319
QY 181 CTGTGACAGCGCTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGGGGGCAACC 240
DB 9320 CTGTGACAGCGCTAGTACCAAGGGCTGCACCCCGGTTTTTGTTCGAAGGGGGCAACC 9379

```
QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 9380 CCCGCTTGGAAATTAATAAACT 9399

RESULT 2
US-10-189-359-2
; Sequence 2, Application US/10189359
; Publication NO. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Annette
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9399
; TYPE: DNA
; ORGANISM: GBV-A-like virus
US-10-189-359-2

Query Match 100.0%; Score 260; DB 7; Length 9399;
Best Local Similarity 100.0%; Pred. No. 6.2e-82;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 9140 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 9199

QY 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
Db 9200 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 9259

QY 121 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTGTGTCAGCGCTCAGCAGTATTTGTCCG 180
Db 9260 TGGGAAGCAGTCAGTATAATCCCGTCGTGTGTGTGTCAGCGCTCAGCAGTATTTGTCCG 9319

QY 181 CTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACC 240
Db 9320 CTGTGCAGCGTAGTACCAAGGCTGCACCCCGTTTGTTCGAAGCGGAGGCAACC 9379

QY 241 CCCGCTTGGAAATTAATAAACT 260
Db 9380 CCCGCTTGGAAATTAATAAACT 9399

RESULT 3
US-10-501-412-1
; Sequence 1, Application US/10501412
; Publication NO. US20050239205A1
; GENERAL INFORMATION:
; APPLICANT: Graziani, Rita
; APPLICANT: Paonessa, Giacomo
; APPLICANT: Tranboni, Cinzia
; TITLE OF INVENTION: GB VIRUS B BASED REPLICONS AND REPLICON
; FILE REFERENCE: ITR0037YP
; CURRENT APPLICATION NUMBER: US/10/501,412
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: PCT/EP03/00281
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,655
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/348,573
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 9397
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GBV-B Replicon
US-10-501-412-2
; Sequence 2, Application US/10501412
; Publication NO. US20050239205A1
; GENERAL INFORMATION:
; APPLICANT: De Tomasi, Amedeo
; APPLICANT: Graziani, Rita
; APPLICANT: Paonessa, Giacomo
; APPLICANT: Tranboni, Cinzia
; TITLE OF INVENTION: GB VIRUS B BASED REPLICONS AND REPLICON
; FILE REFERENCE: ITR0037YP
; CURRENT APPLICATION NUMBER: US/10/501,412
; CURRENT FILING DATE: 2004-07-12
; PRIOR APPLICATION NUMBER: PCT/EP03/00281
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,655
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/348,573
; PRIOR FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 9397
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: GBV-B Replicon
US-10-501-412-2

Query Match 99.4%; Score 258.4; DB 9; Length 9397;
Best Local Similarity 99.6%; Pred. No. 2.3e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 60
Db 9138 GAGTTTGGCGACCATGGTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGAT 9197

QY 61 GACGTCCCTTCTGGCTCATCCACAAAACCGTCTCGGTGGGTGAGGAGTCTCTGGCTGTG 120
```


Db 9198 GACGTCCTCTGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTTGCTGTG 9257
Qy 121 TGGGAAGCAGTCAGTATATCCCGTCGTGTGTGTGACGCTCAGGACGTATTTGTCCG 180
Db 9258 TGGGAAGCAGTCAGTATATCCCGTCGTGTGTGTGACGCTCAGGACGTATTTGTCCG 9317
Qy 181 CTGTCCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 240
Db 9318 CTGTCCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACC 9377
Qy 241 CCGCTTGAATTAATAACT 260
Db 9378 CCGCTTGAATTAATAACT 9397

RESULT 5
US-10-189-359-1
; Sequence 1, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: SANGAR, Annette
; APPLICANT: LEMON, STANLEY M.
; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 260
; TYPE: RNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: GB VIRUS B
US-10-189-359-1

Query Match 97.8%; Score 254.4; DB 7; Length 260;
Best Local Similarity 75.8%; Pred. No. 2.6e-80;
Matches 194; Conservative 61; Mismatches 1; Indels 0; Gaps 0;
Qy 5 TTGGCGACCATGTGATCAGAACCGTTTCGGGTGAAGCCATGTCTGAAGGGGATGACG 64
Db 4 UUGCGACCAUGGUGAUCAGAACCGUUUUGGGUGAAGCUGGUCUGAAGGGGAGUAGCG 63
Qy 65 TCCTTCTGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTTGCTGTGTGGG 124
Db 64 UCCUUCUGGCUUACCAACAAACCGUCUCGGUGGUGAGGAGUCCUGGUGUGGG 123
Qy 125 AAGCAGTCAGTATATCCCGTCGTGTGTGTGACGCTCAGACGTATTTGTCCGCTGT 184
Db 124 AAGCAGUCAGUAUAUCCGUGUGUGUGGACGCCUCACGACGUAUUGUCCGUGU 183
Qy 185 GCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTCCAAAGCGGAGGCAACCCCG 244
Db 184 GCAGAGCGTAGTACCAAGGGCGUGCACCCCGUUUUUUGUCCAAAGCGGAGGCAACCCCG 243
Qy 245 CTGGAATTAATAACT 260
Db 244 CUUGGAUUUAAAAACU 259

RESULT 6
US-10-189-359-12/c
; Sequence 12, Application US/10189359
; Publication No. US20040039187A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Annette
; APPLICANT: SANGAR, DAVID V.
; APPLICANT: LEMON, STANLEY M.

; TITLE OF INVENTION: Chimeric GB Virus B (GBV-B)
; FILE REFERENCE: UTSG:258US
; CURRENT APPLICATION NUMBER: US/10/189,359
; CURRENT FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 10/189,359
; PRIOR FILING DATE: 2002-07-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 108
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer
US-10-189-359-12

Query Match 25.4%; Score 66; DB 7; Length 108;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 195 GTACCAAGGGCTGCACCCCGTTTTTTGTTCCAAAGCGGAGGCAACCCCGCTTGGAAATTA 254
Db 108 GTACCAAGGGCTGCACCCCGTTTTTTGTTCCAAAGCGGAGGCAACCCCGCTTGGAAATTA 49
Qy 255 AAAACT 260
Db 48 AAAACT 43

RESULT 7
US-10-168-651-36/c
; Sequence 36, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, DYUNG Aina M.
; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameen R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT APPLICATION NUMBER: US/10/168,651
; CURRENT FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
; 60/181,625
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PERL Program
; SEQ ID NO 36
; LENGTH: 3179
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 3232992CB1
US-10-168-651-36

Query Match 12.3%; Score 32; DB 6; Length 3179;
Best Local Similarity 50.7%; Pred. No. 1.5;
Matches 77; Conservative 0; Mismatches 75; Indels 0; Gaps 0;


```
QY 225 CAA 227
Db 122 TAA 124

RESULT 11
US-10-000-897-36
; Sequence 36, Application US/10000897
; Publication No. US20030165852A1
; GENERAL INFORMATION:
; APPLICANT: Schueler, P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE
; FILE OF INVENTION: MATERNAL CIRCULATION
; FILE REFERENCE: 11012-004-999
; CURRENT APPLICATION NUMBER: US/10/000,897
; CURRENT FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: 60/248,882
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36:
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-000-897-36

Query Match 11.9%; Score 31; DB 6; Length 1095;
Best Local Similarity 47.6%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;
Matches 91; Conservative 0;

QY 4 TTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGAC 63
Db 620 TTTGCAATGCTCATTAATATTTTCTATTGTTAGACCATATCTGAAGAAATAAC 679

QY 64 GTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTCGCTGTGTGG 123
Db 680 ATTTTAAAGGCTCTACCATAGACATATCTGCTAGAAATGTGTGTGTGTGTGT 739

QY 124 GAAGCGATCAGTATATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
Db 740 GTGTATGTATGTATAGTTCGGGGAGAGGATAGTGTGTGTGTGTGTGTGTGTGTGT 799

QY 184 TGCAGAGCGTA 194
Db 800 GGAGGACTGGA 810

RESULT 12
US-10-818-168-36
; Sequence 36, Application US/10818168
; Publication No. US20040185495A1
; GENERAL INFORMATION:
; APPLICANT: Schueler, P.
; TITLE OF INVENTION: METHODS AND REAGENTS FOR IDENTIFYING RARE FETAL CELLS IN THE MAT
; FILE OF INVENTION: CIRCULATION
; FILE REFERENCE: 11012-004-999
; CURRENT APPLICATION NUMBER: US/10/818,168
; CURRENT FILING DATE: 2004-04-05
; PRIOR APPLICATION NUMBER: 60/248,882
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 36:
; LENGTH: 1095
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-818-168-36

Query Match 11.9%; Score 31; DB 8; Length 1095;
Best Local Similarity 47.6%; Pred. No. 2.6; Mismatches 0; Indels 0; Gaps 0;
Matches 91; Conservative 0;

QY 4 TTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGAC 63
Db 620 TTTGCAATGCTCATTAATATTTTCTATTGTTAGACCATATCTGAAGAAATAAC 679

QY 64 GTCCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGGTGAGAGTCTCGCTGTGTGG 123
Db 680 ATTTTAAAGGCTCTACCATAGACATATCTGCTAGAAATGTGTGTGTGTGTGTGT 739

QY 124 GAAGCGATCAGTATATATCCCTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 183
Db 740 GTGTATGTATGTATAGTTCGGGGAGAGGATAGTGTGTGTGTGTGTGTGTGTGTGT 799

QY 184 TGCAGAGCGTA 194
Db 800 GGAGGACTGGA 810

RESULT 13
US-09-925-065A-551953
; Sequence 551953, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/361,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 551953
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-551953

Query Match 11.8%; Score 30.6; DB 4; Length 1150;
Best Local Similarity 58.1%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 54; Conservative 0;

QY 8 GCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACGTCC 67
Db 998 GTGCACATGGAGCCTTAGCTCTGATTCCAGTTTCCAGTTCAGCAAAATTTATGAAGAAGATGATTTCCC 1057

QY 68 CTTCTGGCTCATCCACAAAACCGTCTCGGGTG 100
Db 1058 CATATGGATGACAGATAAAACACCCCTGTGGCGAG 1090

RESULT 14
US-11-097-143-31040
; Sequence 31040, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; FILE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
```

! PRIOR APPLICATION NUMBER: 60/160,191
! PRIOR FILING DATE: 1999-10-19
! PRIOR APPLICATION NUMBER: 60/161,932
! PRIOR FILING DATE: 1999-10-28
! PRIOR APPLICATION NUMBER: 60/164,769
! PRIOR FILING DATE: 1999-11-12
! PRIOR APPLICATION NUMBER: 60/173,383
! PRIOR FILING DATE: 1999-12-28
! PRIOR APPLICATION NUMBER: 60/175,693
! PRIOR FILING DATE: 2000-01-12
! PRIOR APPLICATION NUMBER: 60/184,831
! PRIOR FILING DATE: 2000-02-24
! PRIOR APPLICATION NUMBER: 60/191,637
! PRIOR FILING DATE: 2000-03-23
! NUMBER OF SEQ ID NOS: 43008
! SOFTWARE: FastSeq for Windows Version 4.0
! SEQ ID NO 31040
! LENGTH: 3772
! TYPE: DNA
! ORGANISM: DROSOPHILA
US-11-097-143-31040

Query Match 11.8%; Score 30.6; DB 10; Length 3772;
Best Local Similarity 58.1%; Pred. No. 4.9;
Matches 54; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 54 AGGGATGACGTCCTCTGGCTCATCCACAAACCCTCTCGGGTGGGTGAGGAGTCCT 113
DB 73 AGCGCTGAAGTCCTGGCTGACTGAACGAACGCCGTTCCGGCACGTGTTGAGTTGG 132
QY 114 GGCTGTGGGAAGCAGTCAGTATATTCCT 146
DB 133 CTCGAGGTATGGCAGTCCTGGAGTTCCCT 165

RESULT 15
US-09-925-065A-425967
! Sequence 425967, Application US/09925065A
! Publication No. US20050228172A9
! GENERAL INFORMATION:
! APPLICANT: Wang, David G.
! TITLE OF INVENTION: Identification and Mapping of Single
! FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
! FILE REFERENCE: 108827.135
! CURRENT APPLICATION NUMBER: US/09/925,065A
! CURRENT FILING DATE: 2001-08-08
! PRIOR APPLICATION NUMBER: US 60/243,096
! PRIOR FILING DATE: 2000-10-24
! PRIOR APPLICATION NUMBER: US 60/252,147
! PRIOR FILING DATE: 2000-11-20
! PRIOR APPLICATION NUMBER: US 60/250,092
! PRIOR FILING DATE: 2000-11-30
! PRIOR APPLICATION NUMBER: US 60/261,766
! PRIOR FILING DATE: 2001-01-16
! PRIOR APPLICATION NUMBER: US 60/289,846
! PRIOR FILING DATE: 2001-05-09
! NUMBER OF SEQ ID NOS: 957086
! SOFTWARE: FastSeq for Windows Version 4.0
! SEQ ID NO 425967
! LENGTH: 645
! TYPE: DNA
! ORGANISM: Homo sapiens
US-09-925-065A-425967

Query Match 11.6%; Score 30.2; DB 4; Length 645;
Best Local Similarity 52.8%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 105 AGGAGTCCTGGGTGGGAAGCAGTCAGTATATTCCTCGTGTGGTGGAGCCCTC 164
DB 2 AGGCCACCTGGAACCTGACCCACCGCTGCCAGTGGAAACACTGCACGAGTGGTTATGCCTC. 61
QY 165 ACGACGTATTGTCCGCTGTGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTTTGTTTC 224

Db 62 GCCAAGTGTGTGGGATCTGTGTGGGGCTTACCATCACCTGCTACTCCCTCTCCCTGTGC 121
QY 225 CAA 227
DB 122 TAA 124
Search completed: December 7, 2005, 20:59:42
Job time : 539 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 7, 2005, 19:43:05 ; Search time 146 Seconds
(without alignments)
665.769 Million cell updates/sec

Title: US-10-009-002-1_COPY_9140_9399
Perfect score: 260
Sequence: 1 Sagtttggcgacctgtgg.....ccccgttggaataaaact 260

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New.*
1: /cgn2_6/prodata/2/pubnpa/US09_NEW_PUB.seq:
2: /cgn2_6/prodata/2/pubnpa/US06_NEW_PUB.seq:
3: /cgn2_6/prodata/2/pubnpa/US07_NEW_PUB.seq:
4: /cgn2_6/prodata/2/pubnpa/US08_NEW_PUB.seq:
5: /cgn2_6/prodata/2/pubnpa/PCT_NEW_PUB.seq:
6: /cgn2_6/prodata/2/pubnpa/US10_NEW_PUB.seq:
7: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq:
8: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq3:
9: /cgn2_6/prodata/2/pubnpa/US11_NEW_PUB.seq3:
10: /cgn2_6/prodata/2/pubnpa/US60_NEW_PUB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	11.2	1723	7	US-11-051-568-28
2	28.6	11.0	1660	6	US-10-750-185-53296
3	28.6	11.0	319608	7	US-11-145-703-1
4	27.4	10.5	38703	7	US-11-052-544-28
5	27.2	10.5	1926	7	US-11-051-568-26
6	27.2	10.5	162537	7	US-11-121-086-59
7	26.6	10.2	169725	7	US-11-121-086-63
8	26.6	10.2	200628	7	US-11-121-086-62
9	26.4	10.2	1683	6	US-10-750-185-58438
10	26.4	10.2	2078	6	US-10-750-185-32595
11	26.4	10.2	2339	7	US-11-102-240-109
12	26.2	10.1	1267	6	US-10-750-185-39628
13	26.2	10.1	1431	6	US-10-750-185-39518
14	26	10.0	600	6	US-10-750-185-2422
15	26	10.0	1032	6	US-10-750-185-24904
16	26	10.0	3223	6	US-10-750-185-46677
17	26	10.0	167891	7	US-11-121-086-14
18	25.8	9.9	339	7	US-11-060-659-5
19	25.8	9.9	340	7	US-11-060-659-6
20	25.8	9.9	1413	6	US-10-750-185-37592
21	25.8	9.9	4539	6	US-10-750-185-56550
22	25.6	9.8	942	6	US-10-750-185-36167
23	25.6	9.8	1699	6	US-10-750-185-56336

C 24	25.6	9.8	2876	6	US-10-750-185-59433
25	25.6	9.8	128552	7	US-11-121-086-1
C 26	25.6	9.8	166639	7	US-11-121-086-52
C 27	25.4	9.8	797	6	US-10-750-185-36568
C 28	25.4	9.8	976	6	US-10-750-185-29663
C 29	25.4	9.8	1064	6	US-10-750-185-45839
C 30	25.4	9.8	1362	6	US-10-750-185-51469
C 31	25.4	9.8	1459	6	US-10-750-185-47911
C 32	25.4	9.8	1576	6	US-10-750-185-55314
C 33	25.4	9.8	1679	6	US-10-750-185-47978
C 34	25.4	9.8	13672	7	US-11-055-035-2
C 35	25.4	9.8	154548	7	US-11-121-086-33
C 36	25.2	9.7	540	6	US-10-750-185-25574
C 37	25.2	9.7	1479	6	US-10-467-9628-50
C 38	25.2	9.7	2253	6	US-10-750-185-49286
C 39	25.2	9.7	7008	6	US-10-821-234-758
C 40	25.2	9.7	169725	7	US-11-121-086-63
C 41	25.2	9.7	200628	7	US-11-121-086-62
C 42	25	9.6	690	6	US-10-750-185-29148
C 43	25	9.6	1516	6	US-10-750-185-53256
C 44	25	9.6	1673	6	US-10-750-185-32424
C 45	25	9.6	2307	6	US-10-750-185-29133

ALIGNMENTS

RESULT 1
US-11-051-568-28
; Sequence 28, Application US/11051568
; Publication No. US20050255141A1
; GENERAL INFORMATION:
; APPLICANT: OPPERMAN, HERMANN
; OZKAYNAK, ENGIN
; KUBERASAMPATH, THANGAVEL
; RUEGER, DAVID C.
; PANG, ROY H.L.
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/11/051,568
; FILING DATE: 04-Feb-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 10/321,799
; FILING DATE: 17-DEC-2002
; APPLICATION NUMBER: US 09/148,925
; FILING DATE: 8-SEP-1998
; APPLICATION NUMBER: US 08/449,699
; FILING DATE: 24-MAY-1995
; APPLICATION NUMBER: US 08/147,023
; FILING DATE: 1-NOV-1993
; APPLICATION NUMBER: US 07/841,646
; FILING DATE: 21-FEB-1992
; APPLICATION NUMBER: US 07/827,052
; FILING DATE: 28-JAN-1992
; APPLICATION NUMBER: US 07/579,865
; FILING DATE: 7-SEP-1990
; APPLICATION NUMBER: US 07/621,849
; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/621,988

```

; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/810,560
; FILING DATE: 20-DEC-1991
; APPLICATION NUMBER: US 07/569,920
; FILING DATE: 20-AUG-1990
; APPLICATION NUMBER: US 07/600,024
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/599,543
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/616,374
; FILING DATE: 21-NOV-1990
; APPLICATION NUMBER: US 07/483,913
; FILING DATE: 22-FEB-1990
; APPLICATION NUMBER: US 07/179,406
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: US 07/232,630
; FILING DATE: 15-AUG-1988
; APPLICATION NUMBER: US 07/315,342
; FILING DATE: 23-FEB-1989
; APPLICATION NUMBER: US 07/660,162
; FILING DATE: 22-FEB-1991
; APPLICATION NUMBER: US 07/422,699
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,613
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,623
; FILING DATE: 17-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DIANA M. STEEL
; REGISTRATION NUMBER: 43,153
; REFERENCE/DOCKET NUMBER: STK-001CP6C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1723 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; TISSUE TYPE: HIPPOCAMPUS
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 490..1696
; OTHER INFORMATION: /function= "OSTROGENIC PROTEIN"
; /product= "HOP2-PP"
; /note= "HOP2 (cDNA)"
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-11-051-568-28

Query Match 11.2%; Score 29; DB 7; Length 1723;
Best Local Similarity 55.4%; Pred. No. 1.4;
Matches 56; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 1 GAGTTGGCCACCATGGTGGATCAGAACCGTTTCCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 807 GAGCTTCGTTACATGGTGGAGGAGACCGTGCCCTGGGCCACGAGGAGCCCATTTGGAA 866

QY 61 GAGGTCCCTTCTGGCTATCCACAAACCGTCTCGGGTGG 101
DB 867 GGAGTTCCGCTTTGACCTGACCCAGATCCCGGCTGGGAGG 907

```

RESULT 2

```

US-10-750-185-53296/c
; Sequence 53296, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
;
; QUERY MATCH
; BEST LOCAL SIMILARITY 11.2%; SCORE 29; DB 7; LENGTH 1723;
; MATCHES 56; CONSERVATIVE 0; MISMATCHES 45; INDELS 0; GAPS 0;
;
; QY 1 GAGTTGGCCACCATGGTGGATCAGAACCGTTTCCGGGTGAAGCCATGGTCTGAAGGGGAT 60
; DB 807 GAGCTTCGTTACATGGTGGAGGAGACCGTGCCCTGGGCCACGAGGAGCCCATTTGGAA 866
;
; QY 61 GAGGTCCCTTCTGGCTATCCACAAACCGTCTCGGGTGG 101
; DB 867 GGAGTTCCGCTTTGACCTGACCCAGATCCCGGCTGGGAGG 907
;
; REMAINING PRIOR APPLICATION DATA REMOVED - SEE FILE WRAPPER OR PALM.
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA

```

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 53296
; LENGTH: 1660
; TYPE: DNA
; ORGANISM: Bovine 19866880574203
US-10-750-185-53296

Query Match 11.0%; Score 28.6; DB 6; Length 1660;
Best Local Similarity 64.2%; Pred. No. 1.9;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 184 TGCAGAGCGTAGTACCAAGGGCTGCACCCCGGTTTTGTTCACGCGGAGGCGAACCCCC 243
DB 84 TGCAGAGCGGAGCAGCTCATGCTACGCGAGATCTAGGTTCAACCCGGAATCCACCACC 25

QY 244 GCTTGA 250
DB 24 TGCTGA 18

```

RESULT 3

```

US-11-145-703-1
; Sequence 1, Application US/11145703
; Publication No. US20050260667A1
; GENERAL INFORMATION:
; APPLICANT: Cohen, Daniel
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Chumakov, Ilya
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Bihain, Bernard
; APPLICANT: Essioux, Laurent
; TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
; FILE REFERENCE: 53-US16.DIV
; CURRENT APPLICATION NUMBER: US/11/145,703
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US/10/147,603
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 09/539,333
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: US 60/126,903
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 60/131,971
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/132,065
; PRIOR FILING DATE: 1999-04-30
; PRIOR APPLICATION NUMBER: US 60/143,928
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 60/145,915
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: US 60/146,453
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/146,452
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: US 60/162,288
; PRIOR FILING DATE: 1999-10-28
; REMAINING PRIOR APPLICATION DATA REMOVED - SEE FILE WRAPPER OR PALM.
; NUMBER OF SEQ ID NOS: 234
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 319608
; TYPE: DNA

```


ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 31..1107
OTHER INFORMATION: 5'regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 1108..1289
OTHER INFORMATION: exon A g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 14877..14920
OTHER INFORMATION: exon B g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 18778..18862
OTHER INFORMATION: exon Bbis g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 25593..25740
OTHER INFORMATION: exon C g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29388..29502
OTHER INFORMATION: exon D g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 29967..30282
OTHER INFORMATION: exon E g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 64666..64812
OTHER INFORMATION: exon F g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 65505..65853
OTHER INFORMATION: exon G g35018 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 65854..67854
OTHER INFORMATION: 3'regulatory region g35018 gene
FEATURE:
NAME/KEY: exon
LOCATION: 94124..94964
OTHER INFORMATION: exon g35017
FEATURE:
NAME/KEY: exon
LOCATION: 201188..201234
OTHER INFORMATION: exon S g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 214676..214793
OTHER INFORMATION: exon T g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215702..215746
OTHER INFORMATION: exon U g35030 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216836..216915
OTHER INFORMATION: exon V g35030 gene
FEATURE:
NAME/KEY: misc feature
LOCATION: 213818..215818
OTHER INFORMATION: 3'regulatory region g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215941
OTHER INFORMATION: exon R complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 215819..215975
OTHER INFORMATION: exon Rbis complement g34872 gene

FEATURE:
NAME/KEY: exon
LOCATION: 216661..216952
OTHER INFORMATION: exon Qbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 216661..217061
OTHER INFORMATION: exon Q complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 217027..217061
OTHER INFORMATION: exon Q1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 229647..229742
OTHER INFORMATION: exon X complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 230408..230721
OTHER INFORMATION: exon P complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231272..231412
OTHER INFORMATION: exon Obis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231787..231880
OTHER INFORMATION: exon O2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 231870..231879
OTHER INFORMATION: exon O1 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 234174..234321
OTHER INFORMATION: exon O complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 237406..237428
OTHER INFORMATION: exon Nbis complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239807
OTHER INFORMATION: exon N2 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 239719..239853
OTHER INFORMATION: exon N complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240569
OTHER INFORMATION: exon M1117 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240596
OTHER INFORMATION: exon M1090 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240617
OTHER INFORMATION: exon M1069 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240644
OTHER INFORMATION: exon M52 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240824
OTHER INFORMATION: exon M862 complement g34872 gene
FEATURE:
NAME/KEY: exon
LOCATION: 240528..240994
OTHER INFORMATION: exon M692 complement g34872 gene
FEATURE:

```
; NAME/KEY: exon
; LOCATION: 240528..241685
; OTHER INFORMATION: exon M1 complement g34872 gene
; FEATURE:
; NAME/KEY: exon
; LOCATION: 240800..240993
; OTHER INFORMATION: exon M51 complement g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 241686..243885
; OTHER INFORMATION: 5'regulatory region g34872 gene
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 290652..292652
; OTHER INFORMATION: 3'regulatory region g34665 gene

Query Match 11.0%; Score 28.6; DB 7; Length 319608;
Best Local Similarity 55.6%; Pred. No. 11;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 35 GGGTGAAGCCATGGTCTGAGGGGATGACGTCCCTTCTGGCTCATCCACAAAACCGTCT 94
Db 21445 GCGTGTTCATGTATGAATAATTAATCAGGTAAATTCAGGTGCTTCAAAATGCTGTCT 21504
QY 95 CGGGTGGGTGAGGAGTCCCTGGCTGTGTGGGAAGCAGTCA 133
Db 21505 CAGGTGTGACTGAAGTGTGTGATGATGGAAAGTACTCA 21543

RESULT 4
US-11-052-544-28
; Sequence 28, Application US/11052544
; Publication No. US20050255504A1
; GENERAL INFORMATION:
; APPLICANT: PARL, Fritz F.
; TITLE OF INVENTION: METHOD OF DETECTING AN INCREASED
; FILE REFERENCE: 22000.0127U2
; CURRENT APPLICATION NUMBER: US/11/052,544
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: 60/543,866
; PRIOR FILING DATE: 2004-02-12
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note =
; OTHER INFORMATION: synthetic construct
US-11-052-544-28

Query Match 10.5%; Score 27.4; DB 7; Length 38703;
Best Local Similarity 50.4%; Pred. No. 14;
Matches 67; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 57 GGATGACGTCCTTCTGGGTCTATCCAAAAACCGTCTCGGGTGGGTGAGGAGTCTCGGC 116
Db 15809 GGATTCCTCCACATGGGGGTTCATCATACAAAAGTTTCTGGGTAGCTGTGAATACTGCC 15868
QY 117 TGTGTGGGAAGCAGTCAGTATATTCCTCGTGTGTGTGAGCTTCAGGAGTATTGTG 176
Db 15869 TGTGGGTTTCAGTATTGAACATCAATGCTACTGTGTGTGCTACTTATTCAGCTGCA 15928
QY 177 TCCGCTGTGCAGA 189
Db 15929 TCCGTGGTCTCTGA 15941

RESULT 5
US-11-051-568-26
; Sequence 26, Application US/11051568
; Publication No. US20050255141A1
; GENERAL INFORMATION:
; APPLICANT: OPPERMANN, HERMANN
; OZKAYNAK, ENGIN
; KUBERASAMPATH, THANGAVEL
; RUEGER, DAVID C.
; PANG, ROY H.L.
; TITLE OF INVENTION: OSTEOGENIC DEVICES
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TESTA, HURWITZ & THIBEAULT
; STREET: 125 HIGH STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: U.S.A.
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/11/051,568
; FILING DATE: 04-Feb-2005
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 10/321,799
; FILING DATE: 17-DEC-2002
; APPLICATION NUMBER: US 09/148,925
; FILING DATE: 8-SEP-1998
; APPLICATION NUMBER: US 08/449,699
; FILING DATE: 24-MAY-1995
; APPLICATION NUMBER: US 08/147,023
; FILING DATE: 1-NOV-1993
; APPLICATION NUMBER: US 07/841,646
; FILING DATE: 21-FEB-1992
; APPLICATION NUMBER: US 07/827,052
; FILING DATE: 28-JAN-1992
; APPLICATION NUMBER: US 07/579,865
; FILING DATE: 7-SEP-1990
; APPLICATION NUMBER: US 07/621,849
; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/621,988
; FILING DATE: 4-DEC-1990
; APPLICATION NUMBER: US 07/810,560
; FILING DATE: 20-DEC-1991
; APPLICATION NUMBER: US 07/569,920
; FILING DATE: 20-AUG-1990
; APPLICATION NUMBER: US 07/600,024
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/599,543
; FILING DATE: 18-OCT-1990
; APPLICATION NUMBER: US 07/616,374
; FILING DATE: 21-NOV-1990
; APPLICATION NUMBER: US 07/483,913
; FILING DATE: 22-FEB-1990
; APPLICATION NUMBER: US 07/179,406
; FILING DATE: 08-APR-1988
; APPLICATION NUMBER: US 07/232,630
; FILING DATE: 15-AUG-1988
; APPLICATION NUMBER: US 07/315,342
; FILING DATE: 23-FEB-1989
; APPLICATION NUMBER: US 07/660,162
; FILING DATE: 22-FEB-1991
; APPLICATION NUMBER: US 07/422,699
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,613
; FILING DATE: 17-OCT-1989
; APPLICATION NUMBER: US 07/422,623
; FILING DATE: 17-OCT-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: DIANA M. STEEL
; REGISTRATION NUMBER: 43,153
```

```
/
/ REFERENCE/DOCKET NUMBER: STK-001CP6C3
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/248-7000
/ TELEFAX: 617/248-7100
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1926 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ ORIGINAL SOURCE:
/ ORGANISM: MURIDAE
/ TISSUE TYPE: EMBRYO
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 93..1289
/ OTHER INFORMATION: /function= "OSTEOGENIC PROTEIN"
/ /product= "mOP2-PP"
/ /notes= "mOP2 CDNA"
/ SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-11-051-568-26

Query Match 10.5%; Score 27.2; DB 7; Length 1926;
Best Local Similarity 53.8%; Pred. No. 6.1; Mismatches 48; Indels 0; Gaps 0;
Matches 56; Conservative 0;

QY 1 GAGTTGGCGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGAT 60
DB 401 GAGCTTCGTCAACATGGTGGAGCGACCGTACCTGGCGCTACGAGGCCACACTGGAA 460

QY 61 GAGCTCCCTTCTGGCTCATCCACAAACCGTCTCGGTGGGTG 104
DB 461 GGAATTCACCTTGGACCTAACCCAGATCCCTGCTGGGGAGGCTG 504

RESULT 6
US-11-121-086-59
; Sequence 59, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 59
; LENGTH: 162537
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-59

Query Match 10.5%; Score 27.2; DB 7; Length 162537;
Best Local Similarity 64.3%; Pred. No. 27; Mismatches 23; Indels 0; Gaps 0;
Matches 41; Conservative 0;

QY 15 TGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACGTCCTCTCTGG 74
DB 51382 TTGAGGACCCCATGCTGATGTGTAACCCCTGCTCTGAGTGGTGGCTTCCTTCAGC 51441

QY 75 CTCA 78
DB 51442 CCA 51445

RESULT 7
US-11-121-086-63
; Sequence 63, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 63
; LENGTH: 169725
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-63

Query Match 10.2%; Score 26.6; DB 7; Length 169725;
Best Local Similarity 60.3%; Pred. No. 44; Mismatches 29; Indels 0; Gaps 0;
Matches 44; Conservative 0;

QY 74 GCTCATCCACAAACCGTCTCGGTGGGTGAGGAGTCTCGTGTGTGGGAGCAGTCA 133
DB 16630 GCTTGGGCACACGACGCGTCTGAAGTGTGTGTGGGCCCTGCCAGCGTGGTTAGCAGACA 16689

QY 134 GTATAATTCCCGT 146
DB 16690 GAGCTGCTGCCGT 16702

RESULT 8
US-11-121-086-62/c
; Sequence 62, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO: 62
; LENGTH: 200628
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-62

Query Match 10.2%; Score 26.6; DB 7; Length 200628;
Best Local Similarity 60.3%; Pred. No. 46; Mismatches 29; Indels 0; Gaps 0;
Matches 44; Conservative 0;

QY 74 GCTCATCCACAAACCGTCTCGGTGGGTGAGGAGTCTCGTGTGTGGGAGCAGTCA 133
DB 64410 GCTTGGGCACACGACGCGTCTGAAGTGTGTGTGGGCCCTGCCAGCGTGGTTAGCAGACA 64351

QY 134 GTATAATTCCCGT 146
DB 64350 GAGCTGCTGCCGT 64338

RESULT 9
US-10-750-185-58438/c
; Sequence 58438, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
```

```
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 58438
; LENGTH: 1683
; TYPE: DNA
; ORGANISM: Bovine 19866880937133
US-10-750-185-58438

Query Match          10.2%; Score 26.4; DB 6; Length 1683;
Best Local Similarity 59.2%; Pred. No. 11;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 110 TCCTGGCTGTGGGAAGCAGTCAGTATTAATTCCTGCTGTGTGTGAGCGCTCACGAC 169
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 TCAGGCGCAGTGGGAAGCAGTGGGTTTCAGTTCTCTGGTCTCTGAGTGGGA 225

QY 170 GTATTGTCCGCTGTG 185
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
224 AAATTCAACAGCTGAG 209

RESULT 10
US-10-750-185-32595
; Sequence 32595, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 32595
; LENGTH: 2078
; TYPE: DNA
; ORGANISM: Bovine 19866880696449
US-10-750-185-32595

Query Match          10.2%; Score 26.4; DB 6; Length 2078;
Best Local Similarity 59.2%; Pred. No. 12;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 43 CCATGTCTGAGGGGATGACGCTCCCTTCTGGCTCATCCACAAAACCGCTCTCGGTGG 102
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1956 CCAGGTCGCGAGGGAGGGCTCTTGAACGTCGACGACCATGCATCCCTCAGGTAGG 2015

QY 103 TGAGGAGTCTCTGGCTG 118
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2016 AGAACAGGTCGCCG 2031

RESULT 11
US-11-102-240-109
; Sequence 109, Application US/11102240
; Publication No. US20050260647A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ANTIBODIES TO POLYPEPTIDES ENCODED BY A NUCLEIC ACID UNDEREXPRESSE
; FILE REFERENCE: P3230R1C106C
; CURRENT APPLICATION NUMBER: US/11/102,240
; CURRENT FILING DATE: 2005-04-08
; PRIOR APPLICATION NUMBER: 10/063662
; PRIOR FILING DATE: 2002-05-07
; PRIOR APPLICATION NUMBER: 10/006867
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/170262
; PRIOR FILING DATE: 199-12-09
; NUMBER OF SEQ ID NOS: 170
; SEQ ID NO 109
; LENGTH: 2339
; TYPE: DNA
; ORGANISM: Homo Sapien
US-11-102-240-109

Query Match          10.2%; Score 26.4; DB 7; Length 2339;
Best Local Similarity 55.4%; Pred. No. 12;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 55 GGGATGACGTCCTTCTGGCTCATCCACAAAACCGTCTCGGGTGGTGAGGAGTCCTG 114
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
768 GTGGCTGAGTGAGATCTGTGTGAACCTCAAAACTCTACCGCTGTGAGTGGGGTTG 827

QY 115 GCTGTGTGGGAAGCAGTCAGTATTAATTTCCCGT 146
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
828 GCCGTGTCTAAGAAGTGATGGCAAGACTTGT 859

RESULT 12
US-10-750-185-39628/c
; Sequence 39628, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 39628
; LENGTH: 1267
; TYPE: DNA
; ORGANISM: Bovine 19866880941444
US-10-750-185-39628

Query Match          10.1%; Score 26.2; DB 6; Length 1267;
Best Local Similarity 63.5%; Pred. No. 12;
Matches 40; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 12 CCATGTGATCAGAACCGTTTCGGGTGAAGCCATGTGTTGAAGGGGATGAGTCCCTTC 71
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1229 CCAGGAAGGAGCAGAGCCCTTTCCCGTGAGGGCAGAGGCTTTTCACATG 1170
```

QY 72 TGG 74
Db 1169 GGG 1167

RESULT 13

US-10-750-185-39518/c
; Sequence 39518, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 39518
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Bovine 19866880359182
US-10-750-185-39518

Query Match 10.1%; Score 26.2; DB 6; Length 1431;
Best Local Similarity 67.3%; Pred. No. 12;
Matches 37; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY 40 AAGCCATGGTCTGAAGGGGATGACGTCCTTCTGGCTCATCCACAAAACCGTCT 94
Db 619 AAGATAATTTCTGATTAAAGATGACTTCCCTGGTGGCTCAGACAGATAAAGAGTCT 565

RESULT 14

US-10-750-185-2422
; Sequence 2422, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 2422
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine MMBT15901
US-10-750-185-2422

Query Match 10.0%; Score 26; DB 6; Length 600;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 65; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY 18 TGGATCAAAACCGTTTCGGGTGAAGCCATGGTCTCTGAAGGGGATGACGTCCTCTTGTGGCTC 77
Db 166 TAGTGATAGCAGGCTCGGGCAAAAGCTGGGCTGAGAGCCTTGGGTCTCCACTCTCTCTCCG 225

QY 78 ATCCACAAAACCGTCTCGGGTGGGTGAGGAGTCTGCTGTGTGGGAAGCAGTCAGTAT 137
Db 226 CTCTGAGAAAAGCTACTATACAGAGCTCTCTCTCTGCTGTGGGCAAGGAGTTAGGAG 285
QY 138 AATTCCTCGTC 147
Db 286 AATCTCCATC 295

RESULT 15

US-10-750-185-24904/c
; Sequence 24904, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIN version 3.1
; SEQ ID NO 24904
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Bovine 19866880791951
US-10-750-185-24904

Query Match 10.0%; Score 26; DB 6; Length 1032;
Best Local Similarity 59.5%; Pred. No. 13;
Matches 44; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 34 CGGTGAAGCCATGGTCTGAAGGGGATGACGTCCTTCTGGCTCATCCACAAAACCGTC 93
Db 317 CTGGTTCTGACAGGGGTTCAAGAGCGGCGTCCACACAGCCCGGCCCATAGTAACCTTG 258
QY 94 TCGGGTGGGTGAGG 107
Db 257 TGGTTGGGTGAGG 244

Search completed: December 7, 2005, 21:02:13
Job time : 148 secs

THIS PAGE BLANK (USPTO)

Result No.	Query			Length	DB	ID	Description
	Score	Match	No.				
C 1	34.8	13.4	657	10	BX964816	BX964816	Reverse s
C 2	34.8	13.4	670	9	AZ962384	AZ962384	2M0231111
C 3	34.8	13.4	818	11	CR479935	CR479935	Reverse s
C 4	33.8	13.0	1092	11	CNS05POX	AL347370	Tetraodon
C 5	33.6	12.9	890	7	C0738666	CO738666	SL1E04C21
C 6	33.2	12.8	571	1	AL872732	AL872732	AL872732
C 7	33.2	12.8	945	5	BQ919900	BQ919900	AGENCOURT
C 8	33.2	12.8	1508	2	B8615183	B8615183	601281124
C 9	33	12.7	1054	2	B6687472	B6687472	602639444
C 10	32.8	12.6	886	7	CV242705	CV242705	WS02515.B
C 11	32.6	12.5	287	1	AV072177	AV072177	AV072177
C 12	32.6	12.5	539	3	BP744242	BP744242	BP744242
C 13	32.6	12.5	561	9	AQ657755	AQ657755	Sheared D
C 14	32.6	12.5	1164	10	AG065034	AG065034	Pan trogl
C 15	32.6	12.5	1527	10	AG118154	AG118154	Pan trogl
C 16	32.4	12.5	949	5	BQ931612	BQ931612	AGENCOURT
C 17	32.2	12.4	251	2	B8425694	B8425694	BB425694
C 18	32.2	12.4	470	2	B8807696	B8807696	BB807696
C 19	32.2	12.4	645	7	CK525779	CK525779	rswfs0.00
C 20	32	12.3	283	2	B8385246	B8385246	BB835246
C 21	32	12.3	483	3	BM631308	BM631308	UI-E-C11-
C 22	32	12.3	645	11	CR840902	CR840902	GROAAA9DH


```

RESULT 2
AZ962384
LOCUS
DEFINITION
  2M023111F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
  clone UUGC2M023111 F, genomic survey sequence.
ACCESSION
AZ962384
VERSION
KEYWORDS
SOURCE
  GSS.
ORGANISM
  Mus musculus (house mouse)
REFERENCE
  Dunn,D., Aoyagi,A., Barber,M., Becorn,T., Duval,B., Hamil,C.,
  Ismail,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
  Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
  Niederhausern,A. and Wright,D. Weiss R.
  Mouse whole genome scaffolding with paired end reads from 10kb
  plasmid inserts
JOURNAL
COMMENT
  Contact: Robert B. Weiss
  University of Utah Genome Center
  Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
  84112, USA
  Tel: 801 585 5606
  Fax: 801 585 7177
  Email: ddunn@genetics.utah.edu
  Insert Length: 10000 Std Error: 0.00
  Plate: 0231 row: L column: 11
  Seq primer: CGTGTAAACGACGCGCCAGT
  Class: plasmid ends
  High quality sequence stop: 670.
FEATURES
  source
    1..670
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /strain="C57BL/6J"
      /db_xref="taxon:10090"
      /clone="UUGC2M023111"
      /sex="Female"
      /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
      /clone_lib="Mouse 10kb plasmid UUGC2M library"
      /notes="Vector: PWD42nv; Purified genomic DNA from M.
      musculus C57BL/6J (female) was obtained from the Jackson
      Laboratory Mouse DNA Resource
      (http://www.jax.org/resources/documents/dnares/). The DNA
      was hydrodynamically sheared by repeated passage through a
      0.005 inch orifice at constant velocity. The sheared DNA
      was blunt end-repaired with T4 DNA polymerase and T4
      polynucleotide kinase. Adaptor oligonucleotides were
      ligated to the blunt ends in high molar excess. The
      adaptor DNA was purified and size-selected for a 9.5 to
      10.5 kb range using preparative agarose gel
      electrophoresis. Vector DNA was prepared from a derivative
      of pWD42 [G14732114|gb|AF129072.1], a copy-number
      inducible derivative of plasmid R1. The vector was ligated
      with adaptors complementary to the insert adaptors and
      purified. The sheared, adaptor mouse DNA was annealed to
      adaptor vector DNA, and transformed into
      chemically-competent E. coli XL10-Gold (Stratagene) cells
      and selected for ampicillin resistance."
ORIGIN
  Query Match 13.4%; Score 34.8; DB 9; Length 670;
  Best Local Similarity 54.8%; Pred. No. 11;
  Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
  QY 134 GTATAATTCGGTGTGGTGGAGCCCTCAGCAGTATTTCGCGCTGCAGAGCGT 193
  Db 351 GTGTGACTGCGCATGTGTGTGTACCCCAACCCCTATGCTCTGCTGCGCAACATGT 410

```

```

QY 194 AGTACCAAGGGCTGCACCCCGTTTTCCTCAAGCGGAGGCAACCCCGCTTGAATT 253
Db 411 AAAATGAAGATAGCCCTCAGGTGTGTACGATGATGGAACACACATGTGGGAAA 470
QY 254 AAAAAC 259
Db 471 GATTAAC 476
RESULT 3
CR247935
LOCUS
DEFINITION
  CR247935 818 bp DNA linear GSS 06-JUL-2004
  Reverse strand read from insert in 5'HPRT insertion targeting and
  chromosome engineering clone MHPN90111, genomic survey sequence.
ACCESSION
CR247935
VERSION
KEYWORDS
  GSS; genome survey sequence; MICER.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
  Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 818)
  Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L.,
  Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y.,
  Rogers,J. and Bradley,A.
  Direct Submision
  Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire,
  CB10 1SA, UK. http://www.sanger.ac.uk/MICER
FEATURES
  source
    1..818
      /organism="Mus musculus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10090"
      /clone="MHPN90111"
      /clone_lib="MHPN"
ORIGIN
  Query Match 13.4%; Score 34.8; DB 11; Length 818;
  Best Local Similarity 54.8%; Pred. No. 12;
  Matches 69; Conservative 0; Mismatches 57; Indels 0; Gaps 0;
  QY 134 GTATAATTCGGTGTGGTGGAGCCCTCAGCAGTATTTCGCGCTGCAGAGCGT 193
  Db 374 GTGTGACTGCGCATGTGTGTGTACCCCAACCCCTATGCTCTGCTGCGCAACATGT 315
  QY 194 AGTACCAAGGGCTGCACCCCGTTTTCCTCAAGCGGAGGCAACCCCGCTTGAATT 253
  Db 314 AAAATGAAGATAGCCCTCAGGTGTGTACGATGATGGAACACACATGTGGGAAA 255
  QY 254 AAAAAC 259
  Db 254 GATTAAC 249
RESULT 4
CNS05POX/c
LOCUS
DEFINITION
  CNS05POX 1092 bp DNA linear GSS 01-SEP-2000
  Tetraodon nigroviridis genome survey sequence T7 end of clone
  045G18 of library A from Tetraodon nigroviridis, genomic survey
  sequence.
ACCESSION
AL347370
VERSION
KEYWORDS
  GSS; genome survey sequence.
SOURCE
  Tetraodon nigroviridis
ORGANISM
  Tetraodon nigroviridis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
  Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
  Tetraodontidae; Tetraodontidae; Tetraodon.
REFERENCE
  1
  Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,

```

Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,P.,
Saurin,W. and Weissenbach,J.
Estimate of human gene number provided by genome-wide analysis
using Tetraodon nigroviridis DNA sequence
Nat. Genet. 25 (2), 235-238 (2000)
10835645
2

Roest Crolius,H., Jallion,O., Dasilva,C., Ozouf-Costaz,C.,
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,P.,
Saurin,W., Bernot,A. and Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Genome Res. 10 (7), 939-949 (2000)
10899143
3 (bases 1 to 1092)

Genoscope.
Direct Submission
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

Location/Qualifiers
1..1092
/organism="Tetraodon nigroviridis"
/mol_type="genomic DNA"
/db_xref="taxon:99883"
/clone="045G18"
/clone_lib="A"
/notes="Genoscope sequence ID : C0AA045BD09C1
end : T"

Query Match 13.0%; Score 33.8; DB 11; Length 1092;
Best Local Similarity 50.0%; Pred. No. 25;
Matches 83; Conservative 0; Mismatches 83; Indels 0; Gaps 0;
QY 35 GGGTGAGGCGATGGTCTGAAGGGGATGAGTCCCTCTGGCTCATCCACAAAACCGTCT 94
Db 654 GAGTGCAGCGCTTCTTTGTAATGGAGAGGTGAACACACGCCAACCTGCGAAGCGCTTCA 595
QY 95 CGGTGGGTGAGGAGTCCGTGGTGTGGGAGACAGTCAGTATATATCCCGTCTGTGTG 154
Db 594 TCGTTGGGCGAGGAGGAGCGCTCTGTGGTAGCAGACAGATTCAGAGCGCTCTGGGCG 535
QY 155 GTGACGCCCTCAGACGATTTTGTCCGCTGTGACAGCGGTAGTACCA 200
Db 534 GNTCAGACAACTGGATTTTTTTATTTCTTTAGCTAATATTGTTTACA 489

RESULT 5
C0738666 890 bp mRNA linear EST 29-JUL-2004
LOCUS SL1E04c21c04f1 squirrel embryo library 1 Spermophilus lateralis
DEFINITION CDNA clone 21c04 5', mRNA sequence.
ACCESSION C0738666
VERSION C0738666.1 GI:50825936
KEYWORDS EST.
SOURCE Spermophilus lateralis (golden-mantled ground squirrel)
ORGANISM Spermophilus lateralis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Scuridae; Xerinae; Marmotini; Spermophilus.
1 (bases 1 to 890)
REFERENCE Williams,D.R., Gracey,A.Y., Martin,S.L., Hughes,M.A., Li,W.,
Rogers,J. and Cossins,A.R.
AUTHORS Microarray analysis of transcriptional changes during hibernation
in the golden mantled ground squirrel, Spermophilus lateralis
Unpublished (2004)
CONTACT: Andrew R. Cossins
COMMENT Laboratory for Environmental Gene Regulation

University of Liverpool
School of Biological Sciences, The Biosciences Building, Crown
Street, Liverpool, United Kingdom, L69 7ZB
Tel: +44(0)151-795-4510
Fax: +44(0)151-795-4431
Email: cossins@liv.ac.uk
Vector has been trimmed from this EST.
Plate: 21 row: c column: 04
Seq primer: pf1c T7 (5'-AATACGACTCACTATAGG-3')
High quality sequence stop: 890.

Location/Qualifiers
1..890
/organism="Spermophilus lateralis"
/mol_type="mRNA"
/db_xref="taxon:76772"
/clone="21c04"
/sex="male and female"
/tissue_type="embryo"
/dev_stage="embryonic"
/lab_host="E.coli Electromax DH10B"
/clone_lib="squirrel embryo library 1"
/note="Vector: pf1c; Site 1: SalI GTCGAG; Site 2: BamHI
CGATCC; Normalized and subtracted cDNA library_prepared
from embryos"

Query Match 12.9%; Score 33.6; DB 7; Length 890;
Best Local Similarity 52.9%; Pred. No. 28;
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
QY 12 CCATGCTGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGAGTCCCTTC 71
Db 572 CCATGCTGCTTGGCTCTCTATCTGCTGAGGTTTCCCGAGTGAGGATGAGATTTTATA 631
QY 72 TGGCTCATCCACAAAACCGTCTCGGTGGTGAGAGTCTCGGTCTGTGCGAAGCAGT 131
Db 632 AATGTGTCAGTCAAAACAGCTTCCACAAGAGTGAAGAGGCGCTCTGCTGTTGGTGAAGGGG 691
QY 132 CAGTATAATTCCCGTC 147
Db 692 CAGTCCATTTCTGTC 707

RESULT 6
AL872732/c 571 bp mRNA linear EST 03-DEC-2003
LOCUS AL872732 XGC-egg Xenopus tropicalis cDNA clone TEgg135n13 5', mRNA
DEFINITION sequence.
ACCESSION AL872732
VERSION AL872732.2 GI:38666145
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 571)
REFERENCE Croning,M.D.R., Ashurst,J.L., Taylor,R., Zorn,A.M. and Rogers,J.
AUTHORS Sanger Xenopus tropicalis EST project 2001
TITLE TROPICALIS SEQUENCE ID: TEgg135n13.plkSP6
JOURNAL Unpublished (2003)
COMMENT On Sep 15, 2002 this sequence version replaced gi:22892997.
Contact: Taylor R
Sanger Institute
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk

This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
cDNA was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the
5' end and NotI at the 3' end.

Vector: pCS107; Site_1: EcoRI; Site_2: NotI
Host: Escherichia coli XL1-blue.
Location/Qualifiers
1. .571
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TE9135n13"
/dev_stage="egg"
/lab_host="Escherichia coli XL1-blue"
/clone_lib="XGC-egg"
/notes="vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA was oligo dt primed from sug of poly A+ RNA from egg. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end"

ORIGIN

Query Match 12.8%; Score 33.2; DB 1; Length 571;
Best Local Similarity 61.6%; Pred. No. 34;
Matches 53; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 60 TGACGTCCTTCGGCTCATCAAAACCGTCTCGGTGGTGAGGAGCTCGTGT 119
|||
Db 383 TGTGGCTCTCTGGCTCATCACATACACTAAAGATTGGCTTTGGAGCCACAGGT 324
|||

Qy 120 GTGGGAAGCAGTCAGTATATATTCCTCCG 145
|||
Db 323 TGGGTTGTGGCAGGATTTTCCAG 298
|||

RESULT 7
BQ919900
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BQ919900 945 bp mRNA linear EST 20-AUG-2002
AGENCOURT_10016818 NCI_CGAP_Mam2 Mus musculus cDNA clone
IMAGE:6488062 5', mRNA sequence.
BQ919900 1 GI:22334586
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muroidea; Muridae; Murinae; Mus.
1 (bases 1 to 945)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL4034 row: a column: 23
High quality sequence stop: 624.
Location/Qualifiers
1. .945
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N-3"
/db_xref="taxon:10090"
/clone="IMAGE:6488062"
/issue_type="tumor, biopsy sample"
/dev_stage="5 months"
/lab_host="DH10B"
/clone_lib="NCI_CGAP Mam2"
/notes="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Library constructed by Life Technologies. Investigator
providing samples: Gilbert Smith. NIH"

FEATURES
source

Query Match 12.8%; Score 33.2; DB 2; Length 1508;
Best Local Similarity 56.4%; Pred. No. 41;
Matches 62; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

Qy 89 CCCTCTCGGTGGGTGAGGAGTCCTGGCTGTGTGGGAAGCAGTCAGTATATTC 148
Cc
Db 1084 CCGCCCAAAATTTTGTGTGTGCCCGGGGGGGGGGGGGGTATATC 1143
Cc

Qy 149 TGTGTGGTGAACCCCTCACGACTATTTGTCCGCTGTGCAGAGCGGTAGTAC 198
Cc
Db 1144 GGGGGGGTTTCTTCCCTGTGGTGTTCCTCTGGTGTGGGGGGGGTACTAC 1193
Cc

ORIGIN

Query Match 12.8%; Score 33.2; DB 5; Length 945;
Best Local Similarity 54.6%; Pred. No. 38;
Matches 65; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

Qy 74 GCTCATCCAAAAACCGTCTCGGTGGGTGAGGAGTCTCGTGTGTGGGAAGCAGTCA 133
|||
Db 598 GCACACCCGCAACACCCGCTTTGATGGTGTGGAGGACTTCAAGAAGAGAGTGC 657
|||

Qy 134 GTATAATTCCTCTGTGTGTGACCGCTCACGAGTATTTGTCCGCTGTGCAGAG 192
|||
Db 658 CAGTCACCTCCCACTCTAAAGCCGAGGCTTTTAGACATTTATCTNTGCTGTG 716
|||

RESULT 8
BE615183
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BE615183 1508 bp mRNA linear EST 24-AUG-2000
601281124F1 NTH_MGC_39 Homo sapiens cDNA clone IMAGE:3623261 5',
mRNA sequence.
BE615183
BE615183.1 GI:9896782
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1508)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM299 row: o column: 06
High quality sequence stop: 307.
Location/Qualifiers
1. .1508
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3623261"
/issue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_39"
/notes="Organ: pancreas; Vector: pOTB7; Site_1: XhoI;
Site_2: EcoRI; cDNA made by oligo-dt priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Library constructed
by Ling Hong in the laboratory of Gerald M. Rubin
(University of California, Berkeley) using ZAP-cDNA
synthesis kit (Stratagene) and Superscript II RT (Life
Technologies)."

FEATURES
source

[illegible]

RESULT 10
CV242705
LOCUS

```

RESULT 11
AV072177/c
LOCUS
DEFINITION
AV072177 Mus musculus stomach C57BL/6J adult Mus musculus cDNA
clone 2200002K22, mRNA sequence.
AV072177
AV072177.1 GI:5192005
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 287)
Carninci,P., Shibata,K., Ozawa,Y., Konno,H., Itoh,M., Aizawa,K.,
Akahira,S., Akiyama,J., Fukuda,S., Fukunishi,Y., Funayama,T.,
Hara,A., Hayatake,N., Hori,F., Ishikawa,T., Itoh,M., Izawa,M.,
Kawa,J., Kikuchi,N., Kojima,Y., Matsuyama,T., Niitsu,H., Oda,H.,
Owa,C., Sato,K., Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y.,
Sugahara,Y., Suzuki,H., Suzuki,H., Tateno,M., Tomaru,Y.,
Tomimaga,N., Watanabe,S., Yagame,M., Yamamura,T., Yokota,T.,
Yoshino,M., Muramatsu,M., Okazaki,Y. and Hayashizaki,Y.
RIKEN Mouse ESTs
Unpublished (1999)
Contact: Chie Owa
Genome Science Laboratory
RIKEN
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-9145
Fax: 81-298-36-9098
Email: genome-res@rtc.riken.go.jp
Thermostabilization and thermoactivation of thermostable enzymes by
trehalose and its application for the synthesis of full length cDNA
{Proc. Natl. Acad. Sci. U.S.A. 95(2):520-524 (1998)}
Transcriptional sequencing: A method for DNA sequencing using RNA
polymerase {Proc. Natl. Acad. Sci. U.S.A. 95(7):3455-3460 (1998)}
Please visit our web site (http://genome.rtc.riken.go.jp) for
further details.
FEATURES
source
Location/Qualifiers
1..287
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="2200002K22"
/sex="male"
/tissue_type="stomach"
/dev_stage="adult"
/clone_lib="Mus musculus stomach C57BL/6J adult"

ORIGIN
Query Match 12.5%; Score 32.6; DB 1; Length 287;
Best Local Similarity 50.0%; Pred. No. 46;
Matches 80; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 9 CGACCATGGTGGATCAGAACCGTTTCGGGTGAAGCCATGGTCTGAAGGGGATGACGTC 68
Db 166 CATAATATCGGTCCAAACACAGTCTGTGGTCCCTTTGTTGCCCTGGGGATGAAGTAC 107

QY 69 TTCTGGCTCATCCACAAAAACCGTCTCGGGTGGGTGAGGAGTCTCTGTGTGGGAAGC 128
Db 106 TCTTTACACATCTTCATANC CGGTCCGAAGTGGGTTCCATGCAGGGTCTGTGGAAAAAGG 47

QY 129 AGTCAGTATATATCCCGTCGTGTGTGTGTGACGCCCTACGA 168
Db 46 TATGGTTATGACTCCCTTTTGGAAAGGCTGGGTCTCTGAAGA 7

RESULT 12
BP744242
LOCUS
DEFINITION
BP744242 Lingulodinium polyedrum cDNA Lingulodinium polyedrum cDNA
clone LP13EST52H08, mRNA sequence.

```

```

FEATURES
  source
    Location/Qualifiers
      1..561
        /organism="Trypanosoma brucei"
        /mol_type="genomic DNA"
        /strain="TREU927/4 GUTat 10.1"
        /db_xref="taxon:5691"
        /clone_lib="Sheared DNA-21P13"
        /clone_lib="Sheared DNA"
        /note="Vector: pUC18; Site 1: SmaI; Constructed at The
        Institute for Genomic Research (TIGR), Rockville, MD.
        Genomic DNA isolated from a cloned population of
        trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
        sheared to give a tight size distribution (approx 2 kb).
        The v + i method used for the library construction is
        described in detail in Smith, H.O. and Venter, J.C.
        (Waking small insert libraries for whole genome shotgun
        sequencing projects. In Genome Sequencing: A Practical
        Approach, eds. M. Vaudin and B. Borell, Oxford University
        Press, 1999)."
```

Query Match 12.5%; Score 32.6; DB 9; Length 561;
 Best Local Similarity 50.3%; Pred. No. 53; Mismatches 0; Gaps 0;
 Matches 80; Conservative 0; Indels 79; Indels 0; Gaps 0;

QY 90 CGTCTCGGTTGGGTGAGGAGTCTCGTGTGTGGGAAGCAGTCAGTATAATTCCTCGT 149
 DB 479 CTTCACCCAGGCTCGCGATCTGCTCTTTGTGAGCGGCTGGGTGCTCGT 420

QY 150 GTGTGTGACGCTCAGCAGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGCTGCA 209
 DB 419 GTAGGCGGCTGGTGTGCGGCTGCTCGTCTCTTTGTATTGTGTCAGTAGCGCCCGAGTA 360

QY 210 CCCCGTTTTTGTTCACAGGAGGCGCAACCCCGCTTG 248
 DB 359 TGTGCGTTTAGTGCTACTGCTCCGCTCCCGGCTTG 321

RESULT 14
 AG065034/c
 LOCUS
 DEFINITION
 Pan troglodytes DNA, clone: PTB-054E17.R, genomic survey sequence.
 AG065034
 ACCESSION
 VERSION
 AG065034.1 GI:16616836
 KEYWORDS
 GSS.
 SOURCE
 Pan troglodytes (chimpanzee)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pan.

REFERENCE
 1
 FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 1164)
 FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS
 Sequencing: M13Rev
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.

COMMENT
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

```

FEATURES
  source
    Location/Qualifiers
      1..1164
        /organism="Pan troglodytes"
        /mol_type="genomic DNA"
        /db_xref="taxon:9598"
        /clone_lib="PTB-054E17.R"
        /sex="male"
        /cell_type="lymphoblast"
        /clone_lib="PTB Chimpanzee Male BAC Library"
```

Query Match 12.5%; Score 32.6; DB 10; Length 1164;
 Best Local Similarity 50.8%; Pred. No. 61; Mismatches 0; Gaps 0;
 Matches 62; Conservative 0; Indels 60; Indels 0; Gaps 0;

QY 89 CCGTCTCGGTTGGGTGAGGAGTCTCGTGTGTGGGAAGCAGTCAGTATAATTCCTCGT 148
 DB 160 CCTCTGGTCTGAGCGGGGGCGGCGGNNNGGCGCGGNAACCCGNTG 101

QY 149 TGTGTGTGACGCTCAGCAGTATTTGTCCGCTGTGCAGAGCGTAGTACCAAGGCTGC 208
 DB 100 TNCNTNNTTCCCTTAACNCNTATTTTCCGCGCCCGGNAACACCGCGCCG 41

QY 209 AC 210
 DB 40 GC 39

RESULT 15
 AG118154
 LOCUS
 DEFINITION
 Pan troglodytes DNA, clone: PTB-126D09.R, genomic survey sequence.
 AG118154
 ACCESSION
 VERSION
 AG118154.1 GI:16738673
 KEYWORDS
 GSS.
 SOURCE
 Pan troglodytes (chimpanzee)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homnidae; Pan.

REFERENCE
 1
 FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 1527)
 FUJIIYAMA, A., HATTORI, M., TOYODA, A., TAYLOR, T.D., YADA, T.,
 TOTOKI, Y., WATANABE, H. and SAKAKI, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS
 Sequencing: M13Rev
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.

COMMENT
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

FEATURES
 source
 Location/Qualifiers
 1..1527
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone_lib="PTB-126D09.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN

Query Match 12.5%; Score 32.6; DB 10; Length 1527;
 Best Local Similarity 74.5%; Pred. No. 64; Mismatches 14; Indels 0; Gaps 0;
 Matches 41; Conservative 0;

Qy 122 GGGAAAGCAGTCAGTATTAATTCCTGTCGTGTGTGTCACGCTTCACGACGTAATTTG 176
 Db 1081 GTGATGCAGTCAGTATTAAGACACGTCGCGGTGATGAGGCTCTACACATATGTG 1135

Search completed: December 7, 2005, 22:24:30
 Job time : 2440 secs